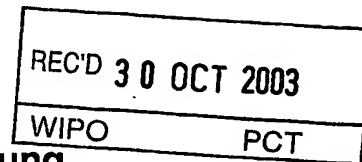




06. 10. 2003



Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 102 39 308.7
Anmeldetag: 27. August 2002
Anmelder/Inhaber: BASF Aktiengesellschaft,
Ludwigshafen/DE
Bezeichnung: Verfahren zur fermentativen Herstellung von
schwefelhaltigen Feinchemikalien
IPC: C 12 P, C 07 C, A 23 K

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 18. September 2003
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Ernst

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Beschreibung

Gegenstand der Erfindung ist ein Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für ein Methylentetrahydrofolat Reduktase (metF)-Gen kodierende Nukleotidsequenzen exprimiert wird.

Stand der Technik

Schwefelhaltige Feinchemikalien, wie zum Beispiel Methionin, Homocystein, S-Adenosyl-Methionin, Glutathion, Cystein, Biotin, Thiamin, Liponsäure werden über natürliche Stoffwechselprozesse in Zellen hergestellt und werden in vielen Industriezweigen verwendet, einschließlich der Nahrungsmittel-, Futtermittel-, Kosmetik- und pharmazeutischen Industrie. Diese Substanzen, die zusammen als "schwefelhaltige Feinchemikalien" bezeichnet werden, umfassen organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Vitamine und Cofaktoren. Ihre Produktion erfolgt am zweckmäßigsten im Großmaßstab mittels Anzucht von Bakterien, die entwickelt wurden, um große Mengen der jeweils gewünschten Substanz zu produzieren und sezernieren. Für diesen Zweck besonders geeignete Organismen sind coryneforme Bakterien, gram-positive nicht-pathogene Bakterien.

Es ist bekannt, dass Aminosäuren durch Fermentation von Stämmen coryneformer Bakterien, insbesondere *Corynebacterium glutamicum*, hergestellt werden. Wegen der großen Bedeutung wird ständig an der Verbesserung der Herstellverfahren gearbeitet. Verfahrensverbesserungen können fermentationstechnische Maßnahmen, wie zum Beispiel Rührung und Versorgung mit Sauerstoff, oder die Zusammensetzung der Nährmedien, wie zum Beispiel die Zuckerkonzentration während der Fermentation, oder die Aufarbeitung zum Produkt, beispielsweise durch Ionenaustauschchromatographie, oder die intrinsischen Leistungseigenschaften des Mikroorganismus selbst betreffen.

Über Stammselektion sind eine Reihe von Mutantenstämmen entwickelt worden, die ein Sortiment wünschenswerter Verbindungen aus der Reihe der schwefelhaltigen Feinchemikalien produzieren. Zur Verbesserung der Leistungseigenschaften dieser Mikroorganismen hinsichtlich der Produktion eines bestimmten Moleküls werden Methoden der Mutagenese, Selektion und Mutantenauswahl angewendet. Dies ist jedoch ein zeitaufwendiges und schwieriges Verfahren. Auf diese Weise erhält man z.B. Stämme, die resistent gegen Antimetabolite, wie z. B. die Methionin-Analoga α -Methyl-Methionin, Ethionin, Norleucin, N-Acetylnorleucin, S-Trifluoromethylhomocystein, 2-Amino-5-heptenolensäure, Seleno-Methionin, Methioninsulfoximin,

Methoxin, 1-Aminocyclopentan-Carboxylsäure oder auxotroph für regulatorisch bedeutsame Metabolite sind und schwefelhaltige Feinchemikalien, wie z. B. L-Methionin, produzieren.

Seit einigen Jahren werden ebenfalls Methoden der rekombinanten DNA-Technik zur Stammverbesserung von L-Aminosäure produzierender Stämme von Corynebacterium eingesetzt, indem man einzelne Aminosäure-Biosynthesegene amplifiziert und die Auswirkung auf die Aminosäure-Produktion untersucht.

Kurze Beschreibung der Erfindung

Der Erfindung lag die Aufgabe zugrunde, ein neues Verfahren zur verbesserten fermentativen Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, bereitzustellen.

Gelöst wird obige Aufgabe durch Bereitstellung eines Verfahrens zur fermentativen Herstellung einer schwefelhaltigen Feinchemikalie, umfassend die Expression einer heterologen Nukleotidsequenz, welche für ein Protein mit metF-Aktivität kodiert, in einem coryneformen Bakterium.

Ein erster Gegenstand der Erfindung ist Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen Feinchemikalie, welches folgende Schritte umfasst:

- a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie produzierenden coryneformen Bakterienkultur, wobei in den coryneformen Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird, welche für ein Protein mit Methylentetrahydrofolat Reduktase (metF)-Aktivität kodiert;
- b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den Zellen der Bakterien, und
- c) Isolieren der schwefelhaltigen Feinchemikalie, welche vorzugsweise L-Methionin umfasst.

Vorzugsweise besitzt obige heterologe metF-kodierende Nukleotidsequenz zur metF-kodierenden Sequenz aus Corynebacterium glutamicum ATCC 13032 eine Sequenzhomologie von weniger als 100%, wie z.B. mehr als 70%, wie 75, 80, 85, 90 oder 95 %, oder weniger als 70%, wie z.B. bis zu 60, 50, 40, 30, 20 oder 10 %. Die metF-kodierende Sequenz ist vorzugsweise aus einem der folgenden Organismen von Liste I abgeleitet:

Liste I

Organismus	Stammsammlung
<i>Corynebacterium diphtheriae</i>	ATCC 14779
<i>Streptomyces lividans</i>	ATCC 19844
<i>Streptomyces coelicolor</i>	ATCC 10147
<i>Aquifex aeolicus</i>	DSM 6858
<i>Burkholderia cepacia</i>	ATCC 25416
<i>Nitrosomonas europaea</i>	ATCC 19718
<i>Pseudomonas aeruginosa</i>	ATCC 17933
<i>Xylella fastidiosa</i>	ATCC 35881
<i>Pseudomonas fluorescens</i>	ATCC 13525
<i>Schizosaccharomyces pombe</i>	ATCC 24969
<i>Saccharomyces cerevisiae</i>	ATCC 10751
<i>Erwinia carotovora</i>	ATCC 15713
<i>Klebsiella pneumoniae</i>	ATCC 700721
<i>Salmonella typhi</i>	ATCC 12839
<i>Salmonella typhimurium</i>	ATCC 15277
<i>Escherichia coli</i> K12	ATCC55151
<i>Vibrio cholerae</i>	ATCC 39315
<i>Haemophilus influenzae</i>	ATCC 51907
<i>Caulobacter crescentus</i>	ATCC 19089
<i>Actinobacillus actinomycetemcomitans</i>	ATCC 33384
<i>Neisseria meningitis</i>	ATCC 6253
<i>Rhodobacter capsulatus</i>	ATCC 11166
<i>Campylobacter jejuni</i>	ATCC 33560
<i>Lactococcus lactis</i>	ATCC 7962
<i>Prochlorococcus marinus</i>	PCC7118
<i>Bacillus stearothermophilus</i>	ATCC 12980

- 5 ATCC: American Type Culture Collection, Rockville, MD, USA
PCC: Pasteur Culture Collection of Cyanobacteria. Paris Frankreich
DSM: Deutsche Sammlung von Mikroorganismen und Zellkulturen

10 Die erfindungsgemäß eingesetzte metF-kodierende Sequenz umfasst vorzugsweise eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit metF-Aktivität kodiert, umfasst.

15 Die erfindungsgemäß eingesetzte metF-kodierende Sequenz kodiert außerdem vorzugsweise für ein Protein mit metF-Aktivität, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52

und 54 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metF-Aktivität steht, umfasst.

- 5 Die kodierende metF-Sequenz ist vorzugsweise eine in coryneformen Bakterien replizierbare oder eine stabil in das Chromosom integrierte DNA oder eine RNA.

Gemäß einer bevorzugten Ausführungsform wird das erfindungsgemäße Verfahren durchgeführt, indem man

- 10 a) einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden metF-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
b) einen Stamm einsetzt, in dem die kodierende metF-Sequenz in das Chromosom des Bakteriums integriert wurde.

- 15 Es ist weiterhin bevorzugt, die kodierende metF-Sequenz für die Fermentation zu überexprimieren.

- 20 Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie verstärkt ist; und / oder
in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet sind, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.

Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie durch Stoffwechselmetabolite in seiner Aktivität nicht in unerwünschter Weise beeinflusst wird.

- 30 Gemäß einer weiteren Ausführungsform des erfindungsgemäßen Verfahrens werden deshalb coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- 35 a) dem für eine Aspartatkinase kodierenden Gen lysC,
b) dem für eine Aspartat-Semialdehyd-Dehydrogenase kodierenden Gen asd
c) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen gap,
d) dem für die 3-Phosphoglycerat Kinase kodierenden Gen pgk,
e) dem für die Pyruvat Carboxylase kodierenden Gen pyc,
f) dem für die Triosephosphat Isomerase kodierenden Gen tpi,

- 5
- g) dem für die Homoserin O-Acetyltransferase kodierenden Gen metA,
 - h) dem für die Cystathionin-gamma-Synthase kodierenden Gen metB,
 - i) dem für die Cystathionin-gamma-Lyase kodierenden Gen metC,
 - j) dem für die Serin-Hydroxymethyltransferase kodierenden Gen glyA,
 - 5 k) dem für die O-Acetylhomoserin-Sulphydrylase kodierenden Gen metY,
 - l) dem für die Methionin Synthase kodierenden Gen meth,
 - m) dem für die Phosphoserin-Aminotransferase kodieren Gen serC
 - n) dem für die Phosphoserin-Phosphatase kodieren Gen serB,
 - o) dem für die Serine Acetyl-Transferase kodieren Gen cysE,
 - 10 p) dem für die Homoserin-Dehydrogenase kodieren Gen hom,
- überexprimiert ist.

15 Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene ausgewählt unter Genen der oben genannten Gruppe a) bis p) mutiert ist, so dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden und dass insbesondere die erfindungsgemäße Produktion der Feinchemikalie nicht beeinträchtigt wird.

20 Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- q) dem für die Homoserine-Kinase kodierenden Gen thrB,
- r) dem für die Threonin Dehydratase kodierenden Gen ilvA,
- s) dem für die Threonin Synthase kodierenden Gen thrC
- t) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen ddh
- u) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen pck,
- v) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen pgj,
- w) dem für die Pyruvat-Oxidase kodierenden Gen poxB,
- x) dem für die Dihydrodipicolinat Synthase kodierenden Gen dapA,
- 30 y) dem für die Dihydrodipicolinat Reduktase kodierenden Gen dapB; oder
- z) dem für die Diaminopicolinat Decarboxylase kodierenden Gen lysA

abschwächt ist, insbesondere durch Verringerung der Expressionsrate des korrespondierenden Gens.

35 Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene der obigen Gruppen

q) bis z) mutiert ist, so dass die enzymatische Aktivität des korrespondierenden Proteins teilweise oder vollständig verringert wird.

5 Vorzugsweise werden in dem erfindungsgemäßen Verfahren Mikroorganismen der Art *Corynebacterium glutamicum* eingesetzt.

Ein weiterer Gegenstand der Erfindung betrifft ein Verfahren zur Herstellung eines L-Methionin-haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst

- 10
- a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
 - b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
 - c) Entfernung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
 - 15 d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.

Gegenstand der Erfindung sind ebenfalls die erstmalig aus obigen Mikroorganismen isolierten kodierenden metF-Sequenzen, die davon kodierten metF-Enzyme sowie die funktionalen Homologen dieser Polynukleotide bzw. Proteine.

20

Detaillierte Beschreibung der Erfindung

a) Allgemeine Begriffe

Als Proteine mit der Aktivität der Methylentetrahydrofolat-Reduktase werden solche Proteine beschrieben, die in der Lage sind 5,10-Methylenetetrahydrofolat ($\text{CH}_2\text{-H(4)Folat}$) unter Oxidation des Cofaktors NADH oder NADPH zu 5-Methyltetrahydrofolat ($\text{CH}_3\text{-H(4)Folat}$) zu reduzieren.

30 Dem Fachmann sind weitere Details des metF Proteins bekannt: (Matthews RG. Sheppard C. Goulding C. European Journal of Pediatrics. 157 Suppl 2:S54-9, 1998, Trimmer EE. Ballou DP. Matthews RG. Biochemistry. 40(21):6205-15, 2001). Der Fachmann kann die enzymatische Aktivität von metF durch Enzymtests nachweisen, Vorschriften dafür können sein: Matthews, R.G., Methylentetrahydrofolate reductase from pig liver. Methods in Enzymology. 122:372-81, 1986.

35

Im Rahmen der vorliegenden Erfindung umfasst der Begriff „schwefelhaltige Feinchemikalie“ jegliche chemische Verbindung, die wenigstens ein Schwefelatom kovalent gebunden enthält

und durch ein erfindungsgemäßes Fermentationsverfahrens zugänglich ist. Nichtlimitierende Beispiele dafür sind Methionin, Homocystein, S-Adenosyl-Methionin, insbesondere Methionin, und S-Adenosyl-Methionin.

- 5 Im Rahmen der vorliegenden Erfindung umfassen die Begriffe „L-Methionin“, „Methionin“, Homocystein und S-Adenosylmethionin auch die korrespondierenden Salze, wie z. B. Methionin-Hydrochlorid oder Methionin-Sulfat.

- 10 "Polynukleotide" bezeichnet im allgemeinen Polyribonukleotide (RNA) und Polydeoxyribonukleotide (DNA), wobei es sich um nicht modifizierte RNA oder DNA oder modifizierte RNA oder DNA handeln kann.

Unter "Polypeptiden" versteht man erfindungsgemäß Peptide oder Proteine, die zwei oder mehr über Peptidbindungen verbundene Aminosäuren enthalten.

15

Der Begriff „Stoffwechselmetabolit“ bezeichnet chemische Verbindungen, die im Stoffwechsel von Organismen als Zwischen- oder auch Endprodukte vorkommen und die neben ihrer Eigenschaft als chemische Bausteine auch modulierende Wirkung auf Enzyme und ihre katalytische Aktivität haben können. Dabei ist aus der Literatur bekannt, dass solche Stoffwechselmetabolite sowohl hemmend als auch stimulierend auf die Aktivität von Enzymen wirken können (Biochemistry, Stryer, Lubert, 1995 W. H. Freeman & Company, New York, New York.). In der Literatur ist auch beschrieben, dass es möglich ist durch Maßnahmen wie Mutation der genomischen DNA durch UV-Strahlung, ionisierender Strahlung oder mutagene Substanzen und nachfolgender Selektion auf bestimmte Phänotypen in Organismen solche Enzyme zu produzieren, in denen die Beeinflussung durch Stoffwechselmetabolite verändert wurde (Sahm H. Eggeling L. de Graaf AA. Biological Chemistry 381(9-10):899-910, 2000; Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94). Diese veränderten Eigenschaften können auch durch gezielte Maßnahmen erreicht werden. Dabei ist dem Fachmann bekannt, dass in Gene für Enzyme auch gezielt bestimmte Nukleotide der für das Protein kodierenden DNA so zu verändern, dass das aus der exprimierten DNA-Sequenz resultierende Protein bestimmte neue Eigenschaften aufweist, so zum Beispiel, dass die modulierende Wirkung von Stoffwechselmetaboliten gegenüber dem nicht veränderten Protein verändert ist

30

Enzyme können derart in ihrer Aktivität beeinflusst werden, dass es zu einer Verringerung der Reaktionsgeschwindigkeit, oder zu einer Veränderung der Affinität gegenüber dem Substrat oder zu einer Änderung der Reaktionsgeschwindigkeiten.

35

Die Begriffe "exprimieren" bzw. "Verstärkung" oder „Überexpression“ beschreiben im Kontext der Erfindung die Produktion bzw. Erhöhung der intrazellulären Aktivität eines oder mehrerer Enzyme in einem Mikroorganismus, die durch die entsprechende DNA kodiert werden. Dazu kann man beispielsweise ein Gen in einen Organismus einbringen, ein vorhandenes Gen durch ein
5 anderes Gen ersetzen, die Kopienzahl des Gens bzw. der Gene erhöhen, einen starken Promotor verwenden oder ein Gen verwenden, das für ein entsprechendes Enzym mit einer hohen Aktivität kodiert und man kann gegebenenfalls diese Maßnahmen kombinieren.

b) Erfindungsgemäße metF-Proteine

10 Erfindungsgemäß mit umfasst sind ebenfalls „funktionale Äquivalente“ der konkret offenbarten metF-Enzyme aus Organismen obiger Liste I.

„Funktionale Äquivalente“ oder Analoga der konkret offenbarten Polypeptide sind im Rahmen
15 der vorliegenden Erfindung davon verschiedene Polypeptide, welche weiterhin die gewünschte biologische Aktivität, wie z.B. Substratspezifität, besitzen.

Unter "funktionalen Äquivalenten" versteht man erfindungsgemäß insbesondere Mutanten, welche in wenigstens einer der oben genannten Sequenzpositionen eine andere als die konkret
20 genannte Aminosäure aufweisen aber trotzdem eine der oben genannten biologische Aktivität besitzen. "Funktionale Äquivalente" umfassen somit die durch eine oder mehrere Aminosäure-Additionen, -Substitutionen, -Deletionen und/oder -Inversionen erhältlichen Mutanten, wobei die genannten Veränderungen in jeglicher Sequenzposition auftreten können, solange sie zu einer Mutante mit dem erfindungsgemäßen Eigenschaftsprofil führen. Funktionale Äquivalenz ist insbesondere auch dann gegeben, wenn die Reaktivitätsmuster zwischen Mutante und unverändertem Polypeptid qualitativ übereinstimmen, d.h. beispielsweise gleiche Substrate mit unterschiedlicher Geschwindigkeit umgesetzt werden.

"Funktionale Äquivalente" umfassen natürlich auch Polypeptide welche aus anderen Organismen zugänglich sind, sowie natürlich vorkommende Varianten. Beispielsweise lassen sich durch
30 Sequenzvergleich Bereiche homologer Sequenzregionen festlegen und in Anlehnung an die konkreten Vorgaben der Erfindung äquivalente Enzyme ermitteln.

„Funktionale Äquivalente“ umfassen ebenfalls Fragmente, vorzugsweise einzelne Domänen oder Sequenzmotive, der erfindungsgemäßen Polypeptide, welche z.B. die gewünschte biologische Funktion aufweisen.
35

„Funktionale Äquivalente“ sind außerdem Fusionsproteine, welche ein der oben genannten Polypeptidsequenzen oder davon abgeleitete funktionale Äquivalente und wenigstens eine weitere, davon funktionell verschiedene, heterologe Sequenz in funktioneller N- oder C-terminaler Verknüpfung (d.h. ohne gegenseitigen wesentliche funktionelle Beeinträchtigung der Fusionsproteinteile) aufweisen. Nichtlimitierende Beispiele für derartige heterologe Sequenzen sind z.B. Signalpeptide, Enzyme, Immunoglobuline, Oberflächenantigene, Rezeptoren oder Rezeptorliganden.

Erfindungsgemäß mit umfasste „funktionale Äquivalente“ sind Homologe zu den konkret offenbarten Proteinen. Diese besitzen wenigstens 30%, oder etwa 40%, 50 %, vorzugsweise wenigstens etwa 60 %, 65%, 70%, oder 75% ins besondere wenigsten 85 %, wie z.B. 90%, 95% oder 99%, Homologie zu einer der konkret offenbarten Sequenzen, berechnet nach dem Algorithmus von Pearson und Lipman, Proc. Natl. Acad. Sci. (USA) 85(8), 1988, 2444-2448.

Homologe der erfindungsgemäßen Proteine oder Polypeptide können durch Mutagenese erzeugt werden, z.B. durch Punktmutation oder Verkürzung des Proteins. Der Begriff „Homolog“, wie er hier verwendet wird, betrifft eine variante Form des Proteins, die als Agonist oder Antagonist der Protein-Aktivität wirkt.

Homologe des erfindungsgemäßen Proteine können durch Screening kombinatorischer Banken von Mutanten, wie z.B. Verkürzungsmutanten, identifiziert werden. Beispielsweise kann eine variierte Bank von Protein-Varianten durch kombinatorische Mutagenese auf Nukleinsäureebene erzeugt werden, wie z.B. durch enzymatisches Ligieren eines Gemisches synthetischer Oligonukleotide. Es gibt eine Vielzahl von Verfahren, die zur Herstellung von Banken potentieller Homologer aus einer degenerierten Oligonukleotidsequenz verwendet werden können. Die chemische Synthese einer degenerierten Gensequenz kann in einem DNA-Syntheseautomaten durchgeführt werden, und das synthetische Gen kann dann in einen geeigneten Expressionsvektor ligiert werden. Die Verwendung eines degenerierten Gensatzes ermöglicht die Bereitstellung sämtlicher Sequenzen in einem Gemisch, die den gewünschten Satz an potentiellen Proteinsequenzen codieren. Verfahren zur Synthese degenerierter Oligonukleotide sind dem Fachmann bekannt (Z.B. Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al. (1983) Nucleic Acids Res. 11:477).

Zusätzlich können Banken von Fragmenten des Protein-Codons verwendet werden, um eine variierte Population von Protein-Fragmenten zum Screening und zur anschließenden Selektion von Homologen eines erfindungsgemäßen Proteins zu erzeugen. Bei einer Ausführungsform kann eine Bank von kodierenden Sequenzfragmenten durch Behandeln eines doppelsträngigen

PCR-Fragmentes einer kodierenden Sequenz mit einer Nuklease unter Bedingungen, unter denen ein Nicking nur etwa einmal pro Molekül erfolgt, Denaturieren der doppelsträngigen DNA, Renaturieren der DNA unter Bildung doppelsträngiger DNA, die Sense-/Antisense-Paare von verschiedenen genickten Produkten umfassen kann, Entfernen einzelsträngiger Abschnitte aus neu gebildeten Duplices durch Behandlung mit S1-Nuclease und Ligieren der resultierenden Fragmentbank in einen Expressionsvektor erzeugt werden. Durch dieses Verfahren kann eine Expressionsbank hergeleitet werden, die N-terminale, C-terminale und interne Fragmente mit verschiedenen Größen des erfindungsgemäßen Proteins kodiert.

Im Stand der Technik sind mehrere Techniken zum Screening von Genprodukten kombinatorischer Banken, die durch Punktmutationen oder Verkürzung hergestellt worden sind, und zum Screening von cDNA-Banken auf Genprodukte mit einer ausgewählten Eigenschaft bekannt. Diese Techniken lassen sich an das schnelle Screening der Genbanken anpassen, die durch kombinatorische Mutagenese von erfindungsgemäßer Homologer erzeugt worden sind. Die am häufigsten verwendeten Techniken zum Screening großer Genbanken, die einer Analyse mit hohem Durchsatz unterliegen, umfassen das Klonieren der Genbank in replizierbare Expressionsvektoren, Transformieren der geeigneten Zellen mit der resultierenden Vektorenbank und Exprimieren der kombinatorischen Gene unter Bedingungen, unter denen der Nachweis der gewünschten Aktivität die Isolation des Vektors, der das Gen codiert, dessen Produkt nachgewiesen wurde, erleichtert. Recursive-Ensemble-Mutagenese (REM), eine Technik, die die Häufigkeit funktioneller Mutanten in den Banken vergrößert, kann in Kombination mit den Screeningtests verwendet werden, um Homologe zu identifizieren (Arkin und Yourvan (1992) PNAS 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331

c) Erfindungsgemäße Polynukleotide

Gegenstand der Erfindung sind ebenso Nukleinsäuresequenzen (einzel- und doppelsträngige DNA- und RNA-Sequenzen, wie z.B. cDNA und mRNA), kodierend für eines der obigen metF-Enzyme und deren funktionalen Äquivalenten, welche z.B. auch unter Verwendung künstlicher Nukleotidanaloga zugänglich sind.

Die Erfindung betrifft sowohl isolierte Nukleinsäuremoleküle, welche für erfindungsgemäße Polypeptide bzw. Proteine oder biologisch aktive Abschnitte davon kodieren, sowie Nukleinsäurefragmente, die z.B. zur Verwendung als Hybridisierungs sonden oder Primer zur Identifizierung oder Amplifizierung von erfindungsgemäßer kodierenden Nukleinsäuren verwendet werden können.

Die erfindungsgemäßen Nukleinsäuremoleküle können zudem untranslatierte Sequenzen vom 3'- und/oder 5'-Ende des kodierenden Genbereichs enthalten

5 Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure zugegen sind und kann überdies im wesentlichen frei von anderem zellulären Material oder Kulturmedium sein, wenn es durch rekombinante Techniken hergestellt wird, oder frei von chemischen Vorstufen oder anderen Chemikalien sein, wenn es chemisch synthetisiert wird.

10 Die Erfindung umfasst weiterhin die zu den konkret beschriebenen Nukleotidsequenzen komplementären Nukleinsäuremoleküle oder einen Abschnitt davon.

15 Die erfindungsgemäß Nukleotidsequenzen ermöglichen die Erzeugung von Sonden und Primern, die zur Identifizierung und/oder Klonierung von homologer Sequenzen in anderen Zelltypen und Organismen verwendbar sind. Solche Sonden bzw. Primer umfassen gewöhnlich einen Nukleotidsequenzbereich, der unter stringenten Bedingungen an mindestens etwa 12, vorzugsweise mindestens etwa 25, wie z.B. etwa 40, 50 oder 75 aufeinanderfolgende Nukleotide eines Sense-Stranges einer erfindungsgemäßen Nukleinsäuresequenz oder eines entsprechenden Antisense-Stranges hybridisiert.

20 Weitere erfindungsgemäße Nukleinsäuresequenzen sind abgeleitet von SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 oder 53 und unterscheiden sich davon durch Addition, Substitution, Insertion oder Deletion einzelner oder mehrerer Nukleotide, kodieren aber weiterhin für Polypeptide mit dem gewünschten Eigenschaftsprofil. Dies können Polynukleotide sein, die zu obigen Sequenzen in mindestens etwa 50%, 55%, 60%, 65%, 70%, 80% oder 90%, vorzugsweise in mindestens etwa 95%, 96%, 97%, 98% oder 99% der Sequenzpositionen identisch sind.

30 Erfindungsgemäß umfasst sind auch solche Nukleinsäuresequenzen, die sogenannte stumme Mutationen umfassen oder entsprechend der Codon-Nutzung eines speziellen Ursprungs- oder Wirtsorganismus, im Vergleich zu einer konkret genannten Sequenz verändert sind, ebenso wie natürlich vorkommende Varianten, wie z.B. Spleißvarianten oder Allelvarianten, davon. Gegenstand sind ebenso durch konservative Nukleotidsubstitutionen (d.h. die betreffende Aminosäure wird durch eine Aminosäure gleicher Ladung, Größe, Polarität und/oder Löslichkeit ersetzt) erhaltliche Sequenzen.

35

Gegenstand der Erfindung sind auch die durch Sequenzpolymorphismen von den konkret offenbarten Nukleinsäuren abgeleiteten Moleküle. Diese genetischen Polymorphismen können zwischen Individuen innerhalb einer Population aufgrund der natürlichen Variation existieren. Diese natürlichen Variationen bewirken üblicherweise eine Varianz von 1 bis 5 % in der Nukleotidsequenz eines Gens.

Weiterhin umfasst die Erfindung auch Nukleinsäuresequenzen, welchen mit oben genannten kodierenden Sequenzen hybridisieren oder dazu komplementär sind. Diese Polynukleotide lassen sich bei Durchmusterung von genomischen oder cDNA-Banken auffinden und gegebenenfalls daraus mit geeigneten Primern mittels PCR vermehren und anschließend beispielsweise mit geeigneten Sonden isolieren. Eine weitere Möglichkeit bietet die Transformation geeigneter Mikroorganismen mit erfindungsgemäßen Polynukleotiden oder Vektoren, die Vermehrung der Mikroorganismen und damit der Polynukleotide und deren anschließende Isolierung. Darüber hinaus können erfindungsgemäße Polynukleotide auch auf chemischem Wege synthetisiert werden.

Unter der Eigenschaft, an Polynukleotide „hybridisieren“ zu können, versteht man die Fähigkeit eines Poly- oder Oligonukleotids unter stringenten Bedingungen an eine nahezu komplementäre Sequenz zu binden, während unter diesen Bedingungen unspezifische Bindungen zwischen nicht-komplementären Partnern unterbleiben. Dazu sollten die Sequenzen zu 70-100%, vorzugsweise zu 90-100%, komplementär sein. Die Eigenschaft komplementärer Sequenzen, spezifisch aneinander binden zu können, macht man sich beispielsweise in der Northern- oder Southern-Blot-Technik oder bei der Primerbindung in PCR oder RT-PCR zunutze. Üblicherweise werden dazu Oligonukleotide ab einer Länge von 30 Basenpaaren eingesetzt. Unter stringenten Bedingungen versteht man beispielsweise in der Northern-Blot-Technik die Verwendung einer 50 – 70 °C, vorzugsweise 60 – 65 °C warmen Waschlösung, beispielsweise 0,1x SSC-Puffer mit 0,1% SDS (20x SSC: 3M NaCl, 0,3M Na-Citrat, pH 7,0) zur Elution unspezifisch hybridisierter cDNA-Sonden oder Oligonukleotide. Dabei bleiben, wie oben erwähnt, nur in hohem Maße komplementäre Nukleinsäuren aneinander gebunden. Die Einstellung stringenter Bedingungen ist dem Fachmann bekannt und ist z.B. in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. beschrieben.

c) Isolierung der kodierenden metF-Gene

Die für das Enzym Methylentetrahydrofolat Reduktase kodierenden metF-Gene aus den Organismen obiger Liste I sind in an sich bekannter Weise isolierbar.

Zur Isolierung der metF-Gene oder auch anderer Gene der Organismen aus obiger Liste I wird zunächst eine Genbank dieses Organismus in Escherichia coli (E. coli) angelegt. Das Anlegen von Genbanken ist in allgemein bekannten Lehrbüchern und Handbüchern ausführlich beschrieben. Als Beispiel seien das Lehrbuch von Winnacker: Gene und Klone, Eine Einführung in die Gentechnologie (Verlag Chemie, Weinheim, Deutschland, 1990), oder das Handbuch von Sambrook et al.: Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1989) genannt. Eine sehr bekannte Genbank ist die des E. coli K-12 Stammes W3110, die von Kohara et al. (Cell 50, 495-508 (198)) in λ -Vektoren angelegt wurde.

Zur Herstellung einer Genbank von Organismen der Liste I in E. coli können Cosmide, wie der Cosmidvektor SuperCos I (Wahl et al., 1987, Proceedings of the National Academy of Sciences USA, 84: 2160-2164), aber auch Plasmide, wie pBR322 (BoliVal; Life Sciences, 25, 807-818 (1979)) oder pUC9 (Vieira et al., 1982, Gene, 19: 259-268), verwendet werden. Als Wirte eignen sich besonders solche E. coli Stämme, die restriktions- und rekombinationsdefekt sind. Ein Beispiel hierfür ist der Stamm DH5 α mcr, der von Grant et al. (Proceedings of the National Academy of Sciences USA, 87 (1990) 4645-4649) beschrieben wurde. Die mit Hilfe von Cosmiden klonierten langen DNA-Fragmente können anschließend wiederum in gängige, für die Sequenzierung geeignete Vektoren subkloniert und anschließend sequenziert werden, so wie es z. B. bei Sanger et al. (proceedings of the National Academy of Sciences of the United States of America, 74: 5463-5467, 1977) beschrieben ist.

Die erhaltenen DNA-Sequenzen können dann mit bekannten Algorithmen bzw. Sequenzanalyse-Programmen, wie z. B. dem von Staden (Nucleic Acids Research 14, 217-232 (1986)), dem von Marck (Nucleic Acids Research 16, 1829-1836 (1988)) oder dem GCG-Programm von Butler (Methods of Biochemical Analysis 39, 74-97 (1998)), untersucht werden.

Die für die metF-Gene kodierenden DNA-Sequenzen von Organismen gemäß obiger Liste I wurde gefunden. Insbesondere wurden DNA-Sequenzen gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53. Weiterhin wurde aus diesen vorliegenden DNA-Sequenzen mit den oben beschriebenen Methoden die Aminosäuresequenzen der entsprechenden Proteine abgeleitet. Durch SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 und 54 sind die sich ergebenden Aminosäuresequenzen der metF-Genprodukte dargestellt.

Kodierende DNA-Sequenzen, die sich aus den Sequenzen gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 durch die Degeneration des genetischen Kodes ergeben, sind ebenfalls Gegenstand der Erfindung. In glei-

cher Weise sind DNA-Sequenzen, die mit diesen Sequenzen oder davon abgeleiteten Sequenzteilen hybridisieren, Gegenstand der Erfindung.

5 Anleitungen zur Identifizierung von DNA-Sequenzen mittels Hybridisierung findet der Fachmann unter anderem im Handbuch "The DIG System Users Guide für Filter Hybridization" der Firma Boehringer Mannheim GmbH (Mannheim, Deutschland, 1993) und bei Liebl et al. (International Journal of Systematic Bacteriology (1991) 41: 255-260). Anleitungen zur Amplifikation von DNA-Sequenzen mit Hilfe der Polymerase-Kettenreaktion (PCR) findet der Fachmann unter anderem im Handbuch von Gait: Oligonukleotide synthesis: A Practical Approach (IRL Press, Oxford, 10 UK, 1984) und bei Newton und Graham: PCR (Spektrum Akademischer Verlag, Heidelberg, Deutschland, 1994).

15 Weiterhin ist bekannt, dass Änderungen am N- und/oder C- Terminus eines Proteins dessen Funktion nicht wesentlich beeinträchtigen oder sogar stabilisieren können. Angaben hierzu findet der Fachmann unter anderem bei Ben-Bassat et al. (Journal of Bacteriology 169: 751-757 (1987)), bei O'Regan et al. (Gene 77: 237-251 (1989), bei Sahin-Toth et al. (Protein Sciences 3: 240-247 (1994)), bei Hochuli et al. (Bionotechnology 6: 1321-1325 (1988)) und in bekannten Lehrbüchern der Genetik und Molekularbiologie.

20 Aminosäuresequenzen, die sich in entsprechender Weise aus den SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 und 54 ergeben, sind ebenfalls Bestandteil der Erfindung.

d) Erfindungsgemäß verwendete Wirtszellen

30 Weitere Gegenstände der Erfindung betreffen als Wirtszelle dienende Mikroorganismen, insbesondere coryneforme Bakterien, die einen Vektor, insbesondere Pendelvektor oder Plasmidvektor, der wenigstens ein metF-Gen erfindungsgemäßer Definition trägt, enthalten oder in denen ein erfindungsgemäßes metF-Gen exprimiert bzw. verstärkt ist.

35 Diese Mikroorganismen können schwefelhaltige Feinchemikalien, insbesondere L-Methionin, aus Glucose, Saccharose, Lactose, Fructose, Maltose, Melasse, Stärke, Cellulose oder aus Glycerin und Ethanol herstellen. Vorzugsweise sind dies coryneforme Bakterien, insbesondere der Gattung Corynebacterium. Aus der Gattung Corynebacterium ist insbesondere die Art Corynebacterium glutamicum zu nennen, die in der Fachwelt für ihre Fähigkeit bekannt ist, L-Aminosäuren zu produzieren.

Als Beispiele für geeignete Stämme coryneformer Bakterien sind solche der Gattung *Corynebacterium*, insbesondere der Art *Corynebacterium glutamicum* (*C. glutamicum*), wie

Corynebacterium glutamicum ATCC 13032,

Corynebacterium acetoglutamicum ATCC 15806,

5 *Corynebacterium acetoacidophilum* ATCC 13870,

Corynebacterium thermoaminogenes FERM BP-1539,

Corynebacterium melassecola ATCC 17965

oder

10 der Gattung *Brevibacterium*, wie

Brevibacterium flavum ATCC 14067

Brevibacterium lactofermentum ATCC 13869 und

Brevibacterium divaricatum ATCC 14020 zu nennen;

oder davon abgeleitete Stämme, wie

15 *Corynebacterium glutamicum* KFCC10065

Corynebacterium glutamicum ATCC21608

welche ebenfalls die gewünschte Feinchemikalie oder deren Vorstufe(n) produzieren.

20 Mit der Abkürzung KFCC ist die Korean Federation of Culture Collection gemeint, mit der Abkürzung ATCC die American type strain culture collection, mit der Abkürzung FERM BP die Sammlung des National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Japan bezeichnet.

e) Durchführung der erfindungsgemäßen Fermentation

Erfindungsgemäß wurde festgestellt, dass coryneforme Bakterien nach Überexpression eines metF-Gens aus Organismen der Liste I in vorteilhafter Weise schwefelhaltige Feinchemikalien, insbesondere L-Methionin, produzieren.

30

Zur Erzielung einer Überexpression kann der Fachmann unterschiedliche Maßnahmen einzeln oder in Kombination ergreifen. So kann die Kopienzahl der entsprechenden Gene erhöht werden, oder es kann die Promotor- und Regulationsregion oder die Ribosomenbindungsstelle, die sich stromaufwärts des Strukturgens befindet, mutiert werden. In gleicher Weise wirken Expressionskassetten, die stromaufwärts des Strukturgens eingebaut werden. Durch induzierbare Promotoren ist es zusätzlich möglich, die Expression im Verlaufe der fermentativen L-Methionin-Produktion zu steigern. Durch Maßnahmen zur Verlängerung der Lebensdauer der mRNA wird

35

ebenfalls die Expression verbessert. Weiterhin wird durch Verhinderung des Abbaus des Enzymproteins ebenfalls die Enzymaktivität verstärkt. Die Gene oder Genkonstrukte können entweder in Plasmiden mit unterschiedlicher Kopienzahl vorliegen oder im Chromosom integriert und amplifiziert sein. Alternativ kann weiterhin eine Überexpression der betreffenden Gene durch Veränderung der Medienzusammensetzung und Kulturführung erreicht werden.

Anleitungen hierzu findet der Fachmann unter anderem bei Martin et al. (Biotechnology 5, 137-146 (1987)), bei Guerrero et al. (Gene 138, 35-41 (1994)), Tsuchiya und Morinaga (Bio/Technology 6, 428-430 (1988)), bei Eikmanns et al. (Gene 102, 93-98 (1991)), in der Europäischen Patentschrift 0472869, im US Patent 4,601,893, bei Schwarzer und Pühler (Biotechnology 9, 84-87 (1991)), bei Remscheid et al. (Applied and Environmental Microbiology 60, 126-132 (1994)), bei LaBarre et al. (Journal of Bacteriology 175, 1001-1007 (1993)), in der Patentanmeldung WO 96/15246, bei Malumbres et al. (Gene 134, 15-24 (1993)), in der japanischen Offenlegungsschrift JP-A-10-229891, bei Jensen und Hammer (Biotechnology and Bioengineering 58, 191-195 (1998)), bei Makrides (Microbiological Reviews 60 : 512-538 (1996)) und in bekannten Lehrbüchern der Genetik und Molekularbiologie.

Gegenstand der Erfindung sind deshalb auch Expressionskonstrukte, enthaltend unter der genetischen Kontrolle regulativer Nukleinsäuresequenzen eine für ein erfindungsgemäßes Polypeptid kodierende Nukleinsäuresequenz; sowie Vektoren, umfassend wenigstens eines dieser Expressionskonstrukte. Vorzugsweise umfassen solche erfindungsgemäßen Konstrukte 5'-stromaufwärts von der jeweiligen kodierenden Sequenz einen Promotor und 3'-stromabwärts eine Terminatorsequenz sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils operativ verknüpft mit der kodierenden Sequenz. Unter einer „operativen Verknüpfung“ versteht man die sequentielle Anordnung von Promotor, kodierender Sequenz, Terminator und gegebenenfalls weiterer regulativer Elemente derart, dass jedes der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann. Beispiele für operativ verknüpfbare Sequenzen sind Aktivierungssequenzen sowie Enhancer und dergleichen. Weitere regulative Elemente umfassen selektierbare Marker, Amplifikationssignale, Replikationsursprünge und dergleichen. Geeignete regulatorische Sequenzen sind z.B. beschrieben in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990).

Zusätzlich zu den artifiziellen Regulationssequenzen kann die natürliche Regulationssequenz vor dem eigentlichen Strukturgen noch vorhanden sein. Durch genetische Veränderung kann diese natürliche Regulation gegebenenfalls ausgeschaltet und die Expression der Gene erhöht oder erniedrigt werden. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das heißt

es werden keine zusätzlichen Regulationssignale vor das Strukturgen insertiert und der natürliche Promotor mit seiner Regulation wird nicht entfernt. Statt dessen wird die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert oder verringert wird. Die Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

Beispiele für brauchbare Promotoren sind: die Promotoren, *ddh*, *amy*, *lysC*, *dapA*, *lysA* aus *Corynebacterium glutamicum*, aber auch gram-positiven Promotoren *SPO2* wie sie in *Bacillus Subtilis* and Its Closest Relatives, Sonenshein, Abraham L., Hoch, James A., Losick, Richard; ASM Press, District of Columbia, Washington und Patek M. Eikmanns B.J. Patek J. Sahm H. Microbiology. 142 1297-309, 1996 beschrieben sind, oder aber auch *cos*-, *tac*-, *trp*-, *tet*-, *trp-tet*-, *lpp*-, *lac*-, *lpp-lac*-, *lacIq*-, *T7*-, *T5*-, *T3*-, *gal*-, *trc*-, *ara*-, *SP6*-, *l-PR*- oder im *l-PL*-Promotor, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden. Bevorzugt ist auch die Verwendung induzierbarer Promotoren, wie z.B. licht- und insbesondere temperaturinduzierbarer Promotoren, wie der *P₁P₁*-Promotor. Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen verwendet werden. Darüber hinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

Die genannten regulatorischen Sequenzen sollen die gezielte Expression der Nukleinsäuresequenzen und der Proteinexpression ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird.

Die regulatorischen Sequenzen bzw. Faktoren können dabei vorzugsweise die Expression positiv beeinflussen und dadurch erhöhen oder erniedrigen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

Die Herstellung einer Expressionskassette erfolgt durch Fusion eines geeigneten Promotors, einer geeigneten Shine-Dalgarno-Sequenz mit einer *metF*-Nukleotidsequenz sowie einem geeigneten Terminationssignal. Dazu verwendet man gängige Rekombinations- und Klonierungstechniken, wie sie beispielsweise in Current Protocols in Molecular Biology, 1993, John Wiley & Sons, Incorporated, New York New York, PCR Methods, Gelfand, David H., Innis, Michael A., Sninsky, John J. 1999, Academic Press, Incorporated, California, San Diego, ., PCR Cloning Protocols, Methods in Molecular Biology Ser., Vol. 192, 2nd ed., Humana Press, New

Jersey, Totowa. T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987) beschrieben sind.

Das rekombinante Nukleinsäurekonstrukt bzw. Genkonstrukt wird zur Expression in einem geeigneten Wirtsorganismus vorteilhafterweise in einen wirtsspezifischen Vektor insertiert, der eine optimale Expression der Gene im Wirt ermöglicht. Vektoren sind dem Fachmann wohl bekannt und können beispielsweise aus "Cloning Vectors" (Pouwels P. H. et al., Hrsg, Elsevier, Amsterdam-New York-Oxford, 1985) entnommen werden. Unter Vektoren sind außer Plasmiden auch alle anderen dem Fachmann bekannten Vektoren, wie beispielsweise Phagen, Transposons, IS-Elemente, Phasmide, Cosmide, und lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können autonom im Wirtsorganismus repliziert oder chromosomal repliziert werden.

Zur Verstärkung wurden erfindungsgemäße metF-Gene beispielhaft mit Hilfe von episomalen Plasmiden überexprimiert. Als Plasmide eignen sich solche, die in coryneformen Bakterien repliziert werden. Zahlreiche bekannte Plasmidvektoren, wie z. B. pZ1 (Menkel et al., Applied and Environmental Microbiology (1989) 64: 549-554), pEKEx1 (Eikmanns et al., Gene 102: 93-98 (1991)) oder pHS2-1 (Sonnen et al., Gene 107: 69-74 (1991)) beruhen auf den kryptischen Plasmiden pHM1519, pBL1 oder pGA1. Andere Plasmidvektoren, wie z. B. pCLIK5MCS, oder solche, die auf pCG4 (US-A 4,489,160) oder pNG2 (Serwold-Davis et al., FEMS Microbiology Letters 66, 119-124 (1990)) oder pAG1 (US-A 5,158,891) beruhen, können in gleicher Weise verwendet werden.

Weiterhin eignen sich auch solche Plasmidvektoren mit Hilfe derer man das Verfahren der Genamplifikation durch Integration in das Chromosom anwenden kann, so wie es beispielsweise von Remscheid et al. (Applied and Environmental Microbiology 60, 126-132 (1994)) zur Duplikation bzw. Amplifikation des hom-thrB-Operons beschrieben wurde. Bei dieser Methode wird das vollständige Gen in einen Plasmidvektor kloniert, der in einem Wirt (typischerweise E. coli), nicht aber in C. glutamicum replizieren kann. Als Vektoren kommen beispielsweise pSUP301 (Simon et al., Bio/ Technology 1, 784-791 (1983)), pK18mob oder pK19mob (Schäfer et al., Gene 145, 69-73 (1994)), Bernard et al., Journal of Molecular Biology, 234: 534-541 (1993)), pEM1 (Schrumpf et al. 1991, Journal of Bacteriology 173: 4510-4516) oder pBGS8 (Spratt et al., 1986, Gene 41: 337-342) in Frage. Der Plasmidvektor, der das zu amplifizierende Gen enthält, wird anschließend durch Transformation in den gewünschten Stamm von C. glutamicum überführt. Methoden zur Transformation sind beispielsweise bei Thierbach et al. (Applied Microbiology and

Biotechnology 29, 356-362 (1988)), Dunican und Shivnan (Biotechnology 7, 1067-1070 (1989)) und Tauch et al. (FEMS Microbiological Letters 123,343-347 (1994)) beschrieben.

Enzyme können durch Mutationen in den korrespondierenden Genen derart in ihrer Aktivität beeinflusst werden, dass es zu einer teilweisen oder vollständigen Verringerung der Reaktionsgeschwindigkeit der enzymatischen Reaktion kommt. Beispiele für solche Mutationen sind dem Fachmann bekannt (Motoyama H. Yano H. Terasaki Y. Anazawa H. Applied & Environmental Microbiology. 67:3064-70, 2001, Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94.)

Zusätzlich kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben einer Expression bzw. Verstärkung eines erfindungsgemäßen metF-Gen eines oder mehrere Enzyme des jeweiligen Biosyntheseweges, des Cystein-Stoffwechselwegs, der Aspartatsemialdehyd-Synthese, der Glykolyse, der Anaplerotik, des Pentose-Phosphat-Stoffwechsels, des Zitronensäure-Zyklus oder des Aminosäure-Exports zu verstärken.

So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, eines oder mehrere der folgenden Gene verstärkt sein:

- das für eine Aspartatkinase kodierende Gen lysC (EP 1 108 790 A2; DNA-SEQ NO. 281),
- das für eine Aspartat-Semialdehyd kodierende Gen asd (EP 1 108 790 A2; DNA-SEQ NO. 282),
- das für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierende Gen gap (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die 3-Phosphoglycerat Kinase kodierende Gen pgk (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die Pyruvat Carboxylase kodierende Gen pyc (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die Triosephosphat Isomerase kodierende Gen tpi (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die Homoserin O-Acetyltransferase kodierende Gen metA (EP 1 108 790 A2; DNA-SEQ NO. 725),
- das für die Cystathionin-gamma-Synthase kodierende Gen metB (EP 1 108 790 A2; DNA-SEQ NO. 3491),
- das für die Cystathionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),
- das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ

NO. 1110),

- das für die O-Acetylhomoserin-Sulfhydrylase kodierende Gen metY (EP 1 108 790 A2; DNA-SEQ NO. 726),

- das für die Methionin Synthase kodierende Gen meth (EP 1 108 790 A2),

5 - das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-SEQ NO. 928)

- eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)

- das für die Serine Acetyl-Transferase kodierende Gen cysE (EP 1 108 790 A2; DNA-SEQ NO. 2818)

10

- das für eine Homoserin-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ NO. 1306)

So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin in coryneformen Bakterien, vorteilhaft sein, gleichzeitig wenigstens eines der nachfolgenden Gene zu mutieren, so dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch einen Stoffwechselmetaboliten in ihrer Aktivität beeinflusst werden:

15

20

- das für eine Aspartatkinase kodierende Gen lysC (EP 1 108 790 A2; DNA-SEQ NO. 281),

- das für die Pyruvat Carboxylase kodierende Gen pyc (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),

- das für die Homoserin O-Acetyltransferase kodierende Gen metA (EP 1 108 790 A2; DNA-SEQ NO. 725),

- das für die Cystathionin-gamma-Synthase kodierende Gen metB (EP 1 108 790 A2; DNA-SEQ NO. 3491),

- das für die Cystathionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),

- das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ NO. 1110),

30

- das für die O-Acetylhomoserin-Sulfhydrylase kodierende Gen metY (EP 1 108 790 A2; DNA-SEQ NO. 726),

- das für die Methionin Synthase kodierende Gen meth (EP 1 108 790 A2),

- das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-SEQ NO. 928)

35

- eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)

- das für die Serine Acetyl-Transferase kodierende Gen *cysE* (EP 1 108 790 A2; DNA-SEQ NO. 2818)
- das für eine Homoserin-Dehydrogenase kodierende Gen *hom* (EP 1 108 790 A2; DNA-SEQ NO. 1306)

5

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsgemäßen *metF*-Gene eines oder mehrere der folgenden Gene abzuschwächen, insbesondere deren Expression zu verringern, oder auszuschalten:

10

- das für die Homoserine-Kinase kodierende Gen *thrB* (EP 1 108 790 A2; DNA-SEQ NO. 3453)
- das für die Threonin Dehydratase kodierende Gen *ilvA* (EP 1 108 790 A2; DNA-SEQ NO. 2328)
- das für die Threonin Synthase kodierende Gen *thrC* (EP 1 108 790 A2; DNA-SEQ NO. 3486)
- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen *ddh* (EP 1 108 790 A2; DNA-SEQ NO. 3494)

15

- das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen *pck* (EP 1 108 790 A2; DNA-SEQ NO. 3157)

20

- das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen *pgi* (EP 1 108 790 A2; DNA-SEQ NO. 950)

- das für die Pyruvat-Oxidase kodierende Gen *poxB* (EP 1 108 790 A2; DNA-SEQ NO. 2873)
- das für die Dihydrodipicolinat Synthase kodierende Gen *dapA* (EP 1 108 790 A2; DNA-SEQ NO. 3476)

- das für die Dihydrodipicolinat Reduktase kodierende Gen *dapB* (EP 1 108 790 A2; DNA-SEQ NO. 3477)

- das für die Diaminopicolinat Decarboxylase kodierende Gen *lysA* (EP 1 108 790 A2; DNA-SEQ NO. 3451)

30

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsgemäßen *metF*-Gene in coryneformen Bakterien gleichzeitig wenigstens eines der folgenden Gene so zu mutieren, dass die enzymatische Aktivität des korrespondierenden Proteins teilweise oder vollständig verringert wird:

35

- das für die Homoserine-Kinase kodierende Gen *thrB* (EP 1 108 790 A2; DNA-SEQ NO. 3453)
- das für die Threonin Dehydratase kodierende Gen *ilvA* (EP 1 108 790 A2; DNA-SEQ NO. 2328)

- das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)
- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2; DNA-SEQ NO. 3494)
- das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)
- das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)
- das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)
- das für die Dihydrodipicolinat Synthase kodierende Gen dapA (EP 1 108 790 A2; DNA-SEQ NO. 3476)
- das für die Dihydrodipicolinat Reduktase kodierende Gen dapB (EP 1 108 790 A2; DNA-SEQ NO. 3477)
- das für die Diaminopicolinat Decarboxylase kodierende Gen lysA (EP 1 108 790 A2; DNA-SEQ NO. 3451)

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben der Expression bzw. Verstärkung eines erfindungsgemäßen metF-Gens unerwünschte Nebenreaktionen auszuschalten (Nakayama: "Breeding of Amino Acid Producing Microorganisms", in: Overproduction of Microbial Products, Krumphanzl, Sikyta, Vaneek (eds.), Academic Press, London, UK, 1982).

Die erfindungsgemäß hergestellten Mikroorganismen können kontinuierlich oder diskontinuierlich im batch- Verfahren (Satzkultivierung) oder im fed batch (Zulaufverfahren) oder repeated fed batch Verfahren (repetitives Zulaufverfahren) zur Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, kultiviert werden. Eine Zusammenfassung über bekannte Kultivierungsmethoden ist im Lehrbuch von Chmiel (Bioprozeßtechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) oder im Lehrbuch von Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)) zu finden.

Das zu verwendende Kulturmedium hat in geeigneter Weise den Ansprüchen der jeweiligen Stämme zu genügen. Beschreibungen von Kulturmedien verschiedener Mikroorganismen sind im Handbuch "Manual of Methods für General Bacteriology" der American Society für Bacteriology (Washington D. C., USA, 1981) enthalten.

Diese erfindungsgemäß einsetzbaren Medien umfassen gewöhnlich eine oder mehreren Kohlenstoffquellen, Stickstoffquellen, anorganische Salze, Vitamine und/oder Spurenelemente.

Bevorzugte Kohlenstoffquellen sind Zucker, wie Mono-, Di- oder Polysaccharide. Sehr gute Kohlenstoffquellen sind beispielsweise Glucose, Fructose, Mannose, Galactose, Ribose, Sorbose, Ribulose, Lactose, Maltose, Saccharose, Raffinose, Stärke oder Cellulose. Man kann Zucker auch über komplexe Verbindungen, wie Melassen, oder andere Nebenprodukte der Zucker-

5 Raffinierung zu den Medien geben. Es kann auch vorteilhaft sein, Gemische verschiedener Kohlenstoffquellen zuzugeben. Andere mögliche Kohlenstoffquellen sind Öle und Fette wie z. B. Sojaöl, Sonnenblumenöl, Erdnußöl und Kokosfett, Fettsäuren wie z. B. Palmitinsäure, Stearinsäure oder Linolsäure, Alkohole wie z. B. Glycerin, Methanol oder Ethanol und organische Säuren wie z. B. Essigsäure oder Milchsäure.

10

Stickstoffquellen sind gewöhnlich organische oder anorganische Stickstoffverbindungen oder Materialien, die diese Verbindungen enthalten. Beispielhafte Stickstoffquellen umfassen Ammoniak-Gas oder Ammoniumsalze, wie Ammoniumsulfat, Ammoniumchlorid, Ammoniumphosphat, Ammoniumcarbonat oder Ammoniumnitrat, Nitrate, Harnstoff, Aminosäuren oder komplexe

15 Stickstoffquellen, wie Maisquellwasser, Sojamehl, Sojaprotein, Hefeextrakt, Fleischextrakt und andere. Die Stickstoffquellen können einzeln oder als Mischung verwendet werden.

20

Anorganische Salzverbindungen, die in den Medien enthalten sein können, umfassen die Chlorid-, Phosphor- oder Sulfatsalze von Calcium, Magnesium, Natrium, Kobalt, Molybdän, Kalium, Mangan, Zink, Kupfer und Eisen

Als Schwefelquelle für die Herstellung von schwefelhaltigen Feinchemikalien, insbesondere von Methionin, können anorganische schwefelhaltige Verbindungen wie beispielsweise Sulfate, Sulfite, Dithionite, Tetrathionate, Thiosulfate, Sulfide aber auch organische Schwefelverbindungen, wie Mercaptane und Thiole, verwendet werden.

Als Phosphorquelle können Phosphorsäure, Kaliumdihydrogenphosphat oder Dikaliumhydrogenphosphat oder die entsprechenden Natrium haltigen Salze verwendet werden.

30

Chelatbildner können zum Medium gegeben werden, um die Metallionen in Lösung zu halten. Besonders geeignete Chelatbildner umfassen Dihydroxyphenole, wie Catechol oder Protocatechuat, oder organische Säuren, wie Citronensäure.

35

Die erfindungsgemäß eingesetzten Fermentationsmedien enthalten üblicherweise auch andere Wachstumsfaktoren, wie Vitamine oder Wachstumsförderer, zu denen beispielsweise Biotin, Riboflavin, Thiamin, Folsäure, Nikotinsäure, Panthothenat und Pyridoxin gehören. Wachstumsfaktoren und Salze stammen häufig von komplexen Medienkomponenten, wie Hefeextrakt, Me-

lassen, Maisquellwasser und dergleichen. Dem Kulturmedium können überdies geeignete Vorstufen zugesetzt werden. Die genaue Zusammensetzung der Medienverbindungen hängt stark vom jeweiligen Experiment ab und wird für jeden spezifischen Fall individuell entschieden. Information über die Medienoptimierung ist erhältlich aus dem Lehrbuch "Applied Microbiol. Physiology, A Practical Approach" (Hrsg. P.M. Rhodes, P.F. Stanbury, IRL Press (1997) S. 53-73, ISBN 0 19 963577 3). Wachstumsmedien lassen sich auch von kommerziellen Anbietern beziehen, wie Standard 1 (Merck) oder BHI (Brain heart infusion, DIFCO) und dergleichen.

Sämtliche Medienkomponenten werden, entweder durch Hitze (20 min bei 1,5 bar und 121°C) oder durch Sterilfiltration, sterilisiert. Die Komponenten können entweder zusammen oder nötigenfalls getrennt sterilisiert werden. Sämtliche Medienkomponenten können zu Beginn der Anzucht zugegen sein oder wahlfrei kontinuierlich oder chargenweise hinzugegeben werden.

Die Temperatur der Kultur liegt normalerweise zwischen 15°C und 45°C, vorzugsweise bei 25°C bis 40°C und kann während des Experimentes konstant gehalten oder verändert werden. Der pH-Wert des Mediums sollte im Bereich von 5 bis 8,5, vorzugsweise um 7,0 liegen. Der pH-Wert für die Anzucht läßt sich während der Anzucht durch Zugabe von basische Verbindungen wie Natriumhydroxid, Kaliumhydroxid, Ammoniak bzw. Ammoniakwasser oder saure Verbindungen wie Phosphorsäure oder Schwefelsäure kontrollieren. Zur Kontrolle der Schaumentwicklung können Antischaummittel wie z. B. Fettsäurepolyglykolester, eingesetzt werden. Zur Aufrechterhaltung der Stabilität von Plasmiden können dem Medium geeignete selektiv wirkende Stoffe, wie z. B. Antibiotika, hinzugefügt werden. Um aerobe Bedingungen aufrechtzuerhalten, werden Sauerstoff oder Sauerstoff haltige Gasmischungen, wie z. B. Umgebungsluft, in die Kultur eingetragen. Die Temperatur der Kultur liegt normalerweise bei 20°C bis 45°C und. Die Kultur wird solange fortgesetzt, bis sich ein Maximum des gewünschten Produktes gebildet hat. Dieses Ziel wird normalerweise innerhalb von 10 Stunden bis 160 Stunden erreicht.

Die so erhaltenen, insbesondere L-Methionin enthaltenden, Fermentationsbrühen haben üblicherweise eine Trockenmasse von 7,5 bis 25 Gew.-%.

Vorteilhaft ist außerdem auch, wenn die Fermentation zumindest am Ende, insbesondere jedoch über mindestens 30% der Fermentationsdauer zuckerlimitiert gefahren wird. Das heißt, dass während dieser Zeit die Konzentration an verwertbarem Zucker im Fermentationsmedium auf ≥ 0 bis 3 g/l gehalten, beziehungsweise abgesenkt wird.

Die Fermentationsbrühe wird anschließend weiterverarbeitet. Je nach Anforderung kann die Biomasse ganz oder teilweise durch Separationsmethoden, wie z. B. Zentrifugation, Filtration,

Dekantieren oder einer Kombination dieser Methoden aus der Fermentationsbrühe entfernt oder vollständig in ihr belassen werden.

Anschließend kann die Fermentationsbrühe mit bekannten Methoden, wie z. B. mit Hilfe eines Rotationsverdampfers, Dünnschichtverdampfers, Fallfilmverdampfers, durch Umkehrosmose, oder durch Nanofiltration, eingedickt beziehungsweise aufkonzentriert werden. Diese aufkonzentrierte Fermentationsbrühe kann anschließend durch Gefriertrocknung, Sprühtrocknung, Sprühgranulation oder durch anderweitige Verfahren aufgearbeitet werden.

Es ist aber auch möglich die schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, weiter aufzureinigen. Hierzu wird die produktthaltige Brühe nach dem Abtrennen der Biomasse einer Chromatographie mit einem geeigneten Harz unterworfen, wobei das gewünschte Produkt oder die Verunreinigungen ganz oder teilweise auf dem Chromatographieharz zurückgehalten werden. Diese Chromatographieschritte können nötigenfalls wiederholt werden, wobei die gleichen oder andere Chromatographieharze verwendet werden. Der Fachmann ist in der Auswahl der geeigneten Chromatographieharze und ihrer wirksamsten Anwendung bewandert. Das gereinigte Produkt kann durch Filtration oder Ultrafiltration konzentriert und bei einer Temperatur aufbewahrt werden, bei der die Stabilität des Produktes maximal ist.

Die Identität und Reinheit der isolierten Verbindung(en) kann durch Techniken des Standes der Technik bestimmt werden. Diese umfassen Hochleistungs-Flüssigkeitschromatographie (HPLC), spektroskopische Verfahren, Färbeverfahren, Dünnschichtchromatographie, NIRS, Enzymtest oder mikrobiologische Tests. Diese Analyseverfahren sind zusammengefaßt in: Patek et al. (1994) Appl. Environ. Microbiol. 60:133-140; Malakhova et al. (1996) Biotekhnologiya 11 27-32; und Schmidt et al. (1998) Bioprocess Engineer. 19:67-70. Ulmann's Encyclopedia of Industrial Chemistry (1996) Bd. A27, VCH: Weinheim, S. 89-90, S. 521-540, S. 540-547, S. 559-566, 575-581 und S. 581-587; Michal, G (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. et al. (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17.

Die Erfindung wird nun anhand der folgenden nicht-limitierenden Beispiele näher beschrieben:

Beispiel 1: Konstruktion von pCLiK5MCS

Zunächst wurden Ampicillinresistenz und Replikationsursprung des Vektors pBR322 mit den Oligonukleotiden p1.3 (SEQ ID NO:55) und p2.3 (SEQ ID NO:56) mit Hilfe der Polymerase-Kettenreaktion (PCR) amplifiziert.

p1.3 (SEQ ID NO:55)

5'-CCCGGGATCCGCTAGCGGCGCGCCGGCCGGCCCGGTGTGAAATACCGCACAG-3'

5 p2.3 (SEQ ID NO:56)

5'-TCTAGACTCGAGCGGCCGCGGCCGGCCTTTAAATTGAAGACGAAAGGGCCTCG-3'

10 Neben den zu pBR322 komplementären Sequenzen, enthält das Oligonukleotid p1.3 (SEQ ID NO:55) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen SmaI, BamHI, NheI und AscI und das Oligonukleotid p2.3 (SEQ ID NO:56) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen XbaI, XhoI, NotI und DraI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,1 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Die stumpfen Enden des DNA-Fragmentes wurden mit dem Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers miteinander ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK1.

Ausgehend vom Plasmid pWLT1 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden neo1 (SEQ ID NO:57) und neo2 (SEQ ID NO:58) eine Kanamycin-Resistenzcassette amplifiziert.

30

neo1 (SEQ ID NO:57):

5'-GAGATCTAGACCCGGGGATCCGCTAGCGGGCTGCTAAAGGAAGCGGA-3'

neo2 (SEQ ID NO:58):

35 5'-GAGAGGCGCGCCGCTAGCGTGGGCGAAGAACTCCAGCA-3'

Neben den zu pWLT1 komplementären Sequenzen, enthält das Oligonukleotid neo1 in 5'-3'

Richtung die Schnittstellen für die Restriktionsendonukleasen XbaI, SmaI, BamHI, NheI und das Oligonukleotid neo2 (SEQ ID NO:58) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen AscI und NheI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,3 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit den Restriktionsendonukleasen XbaI und AscI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK1 wurde ebenfalls mit den Restriktionsendonukleasen XbaI und AscI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,1kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) und Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK2.

Der Vektor pCLiK2 wurde mit der Restriktionsendonuklease DraI (New England Biolabs, Beverly, USA) geschnitten. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde ein ca. 2,3 kb großes Vektorfragment mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers religiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben (1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK3.

- 5 Ausgehend vom Plasmid pWLQ2 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden cg1 ((SEQ ID NO:59) und cg2 (SEQ ID NO:60) der Replikationsursprung pHM1519 amplifiziert.

cg1 (SEQ ID NO:59):

10 5'-GAGAGGGCGGCCGCGCAAAGTCCCGCTTCGTGAA-3'

cg2 (SEQ ID NO:60):

5'-GAGAGGGCGGCCGCTCAAGTCGGTCAAGCCACGC-3'

- 15 Neben den zu pWLQ2 komplementären Sequenzen, enthalten die Oligonukleotide cg1 (SEQ ID NO:59) und cg2 (SEQ ID NO:60) Schnittstellen für die Restriktionsendonuklease NotI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,7 kb wurde mit dem
- 20 GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease NotI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK3 wurde ebenfalls mit der Restriktionsendonuklease NotI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,3kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach An-
- 30 gaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

35

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so

erhaltene Plasmid erhält den Namen pCLiK5.

Für die Erweiterung von pCLiK5 um eine „multiple cloning site“ (MCS) wurden die beide synthetischen, weitestgehend komplementären Oligonukleotide HS445 ((SEQ ID NO:61) und HS446 (SEQ ID NO:62), die Schnittstellen für die Restriktionsendonukleasen SwaI, XhoI, AatI, ApaI, Asp718, MluI, NdeI, SpeI, EcoRV, Sall, ClaI, BamHI, XbaI und SmaI enthalten, durch gemeinsames Erhitzen auf 95°C und langsames Abkühlen zu einem doppelsträngigen DNA-Fragment vereinigt.

10 HS445 (SEQ ID NO:61):

5'-TCGAATTTAAATCTCGAGAGGCCTGACGTCGGGCCCCGGTACCACGCGTCATATGACTAG
TTCGGACCTAGGGATATCGTTCGACATCGATGCTCTTCTGCGTTAATTAACAATTGGGATCC
TCTAGACCCGGGATTAAAT-3'

15 HS446 (SEQ ID NO:62):

5'-GATCATTTAAATCCCGGGTCTAGAGGATCCCAATTGTTAATTAACGCAGAAGAGCATCGA
TGTCGACGATATCCCTAGGTCCGAAGTAGTCATATGACGCGTGGTACCGGGCCCGACGTC
AGGCCTCTCGAGATTAAAT-3'

20 Der Vektor pCLiK5 wurde mit den Restriktionsendonuklease XhoI und BamHI (New England Biolabs, Beverly, USA) geschnitten und mit alkalischer Phosphatase (I (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 5,0 kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem synthetischen doppelsträngigen DNA-Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

35 Die Plasmid-DNA eines individuellen Klon wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5MCS.

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Aca-

demy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pCLiK5MCS ist als SEQ ID NO: 65 aufgeführt.

5

Beispiel 2: Konstruktion von pCLiK5MCS integrativ sacB

Ausgehend vom Plasmid pK19mob (Schäfer et al., Gene 145,69-73(1994)) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden BK1732 und BK1733 das *Bacillus subtilis* sacB Gen (kodierend für Levan Sucrase) amplifiziert.

10

BK1732 (SEQ ID NO:63):

5'-GAGAGCGGCCGCCGATCCTTTTAAACCCATCAC-3'

15 BK1733 (SEQ ID NO:64):

5'-AGGAGCGGCCGCCATCGGCATTTTCTTTTGCG-3'

Neben den zu pEK19mobsac komplementären Sequenzen, enthalten die Oligonukleotide BK1732 und BK1733 Schnittstellen für die Restriktionsendonuklease NotI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,9 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease NotI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

20

Der Vektor pCLiK5MCS (hergestellt gemäß Beispiel 1) wurde ebenfalls mit der Restriktionsendonuklease NotI geschnitten und mit alkalischer Phosphatase (I (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde ein ungefähr 2,4 kb großes Vektorfragment mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente *E.coli* XL-1Blue (Stratagene,

30

35

La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

5 Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5MCS integrativ sacB.

10 Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pCLiK5MCS integrativ sacB ist als SEQ ID NO: 66 aufgeführt. Weitere Vektoren die zur erfindungsgemäßen Expression oder Überproduktion von metF-Genen geeignet sind, können in analoger Weise hergestellt werden.

15

Patentansprüche

1. Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen
Feinchemikalie, welches folgende Schritte umfasst:
- a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie
produzierenden coryneformen Bakterienkultur, wobei in den coryneformen
Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird,
welche für ein Protein mit Methionin-Synthase (metF) –Aktivität kodiert;
- b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den
Zellen der Bakterien, und
- c) Isolieren der schwefelhaltigen Feinchemikalie.
2. Verfahren nach Anspruch 1, wobei die schwefelhaltige Feinchemikalie L-Methionin
umfasst.
3. Verfahren nach einem der vorhergehenden Ansprüche, wobei sich die heterologe metF-
kodierende Nukleotidsequenz zur metF-kodierenden Sequenz aus *Corynebacterium*
glutamicum ATCC 13032 eine Sequenzhomologie vom weniger als 100% aufweist.
4. Verfahren nach Anspruch 3, wobei die metF-kodierende Sequenz aus einem der
folgenden Organismen abgeleitet ist:

Organismus	Stammsammlung
<i>Corynebacterium diptheriae</i>	ATCC 14779
<i>Streptomyces lividans</i>	ATCC 19844
<i>Streptomyces coelicolor</i>	ATCC 10147
<i>Aquifex aeolicus</i>	DSM 6858
<i>Burkholderia cepacia</i>	ATCC 25416
<i>Nitrosomonas europaea</i>	ATCC 19718
<i>Pseudomonas aeruginosa</i>	ATCC 17933
<i>Xylella fastidiosa</i>	ATCC 35881
<i>Pseudomonas fluorescens</i>	ATCC 13525
<i>Schizosaccharomyces pombe</i>	ATCC 24969
<i>Saccharomyces cerevisiae</i>	ATCC 10751
<i>Erwinia carotovora</i>	ATCC 15713
<i>Klebsiella pneumoniae</i>	ATCC 700721
<i>Salmonella typhi</i>	ATCC 12839
<i>Salmonella typhimurium</i>	ATCC 15277
<i>Escherichia coli</i> K12	ATCC55151

<i>Vibrio cholerae</i>	ATCC 39315
<i>Haemophilus influenzae</i>	ATCC 51907
<i>Caulobacter crescentus</i>	ATCC 19089
<i>Actinobacillus actinomycetemcomitans</i>	ATCC 33384
<i>Neisseria meningitis</i>	ATCC 6253
<i>Rhodobacter capsulatus</i>	ATCC 11166
<i>Campylobacter jejuni</i>	ATCC 33560
<i>Lactococcus lactis</i>	ATCC 7962
<i>Prochlorococcus marinus</i>	PCC7118
<i>Bacillus stearothermophilus</i>	ATCC 12980

5. Verfahren nach einem der vorhergehenden Ansprüche, wobei die metF-kodierende Sequenz eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit metF-Aktivität kodiert, umfasst.
6. Verfahren nach einem der vorhergehenden Ansprüche, wobei die metF-kodierende Sequenz für ein Protein mit metF-Aktivität kodiert, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 und 54 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metF-Aktivität steht, umfasst.
7. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metF-Sequenz eine in coryneformen Bakterien replizierbare oder eine stabil in das Chromosom integrierte DNA oder eine RNA ist.
8. Verfahren gemäß Anspruch 7, wobei man
- a) einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden metF-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
- b) einen Stamm einsetzt, in dem die kodierende metF-Sequenz in das Chromosom des Bakteriums integriert wurde
9. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metF-Sequenz überexprimiert wird.

10. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie verstärkt ist oder derart mutiert ist, dass es durch Stoffwechselmetabolite nicht in seiner Aktivität beeinflusst wird.

11. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet sind, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.

12. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- a) dem für eine Aspartatkinase kodierenden Gen *lysC*,
- b) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen *gap*,
- c) dem für die 3-Phosphoglycerat Kinase kodierenden Gen *pgk*,
- d) dem für die Pyruvat Carboxylase kodierenden Gen *pyc*,
- e) dem für die Triosephosphat Isomerase kodierenden Gen *tpi*,
- f) dem für die Homoserin O-Acetyltransferase kodierenden Gen *metA*,
- g) dem für die Cystathionin-gamma-Synthase kodierenden Gen *metB*,
- h) dem für die Cystathionin-gamma-Lyase kodierenden Gen *metC*,
- i) dem für die Serin-Hydroxymethyltransferase kodierenden Gen *glyA*,
- j) dem für die O-Acetylhomoserin-Sulphydrylase kodierenden Gen *metY*,
- k) dem für das *metH* Gen, das für die Vitamin B12 abhängige Methionin-Synthase kodiert,
- l) dem für das *serC* Gen, das für die Phosphoserin-Aminotransferase kodiert,
- m) dem *serB* Gen, das für die Phosphoserin-Phosphatase kodiert,
- n) dem *cysE* Gen, das für die Serine Acetyl-Transferase kodiert, und
- o) dem *hom* Gen, das eine Homoserin-Dehydrogenase kodiert,

überexprimiert oder so mutiert ist, dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden.

13. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneformen Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter
- a) dem für die Homoserine-Kinase kodierenden Gen thrB,
 - b) dem für die Threonin Dehydratase kodierenden Gen ilvA,
 - c) dem für die Threonin Synthase kodierenden Gen thrC
 - d) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen ddh
 - e) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen pck,
 - f) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen pgj,
 - g) dem für die Pyruvat-Oxidase kodierenden Gen poxB,
 - h) dem für die Dihydrodipicolinat Synthase kodierenden Gen dapA,
 - i) dem für die Dihydrodipicolinat Reduktase kodierenden Gen dapB; oder
 - j) dem für die Diaminopicolinat Decarboxylase kodierenden Gen

durch Veränderung der Expressionsrate oder durch Einführung einer gezielten Mutation abgeschwächt ist.

14. Verfahren gemäß einem oder mehreren der vorhergehenden Ansprüche, wobei man Mikroorganismen der Art *Corynebacterium glutamicum* einsetzt.

15. Verfahren zur Herstellung eines L-Methionin haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst

- a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
- b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
- c) Entfernung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
- d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.

16. Verfahren gemäß Anspruch 15, wobei man Mikroorganismen gemäß der Definition in einem der Ansprüche 1 bis 14 einsetzt.

Zusammenfassung

Die Erfindung betrifft Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für

5 ein Methionin-Synthase (metF)-Gen kodierende Nukleotidsequenzen exprimiert wird.

SEQUENZPROTOKOLL

<110> BASF Aktiengesellschaft

<120> MetF

<130> M/43126

<140>

<141>

<160> 66

<210> 1

<211> 984

<212> DNA

<213> corynebacterium diphtheriae

<220>

<221> CDS

<222> (1)..(981)

<223> RDI01260

<400> 1

atg tct gca caa ccg cta cct gct gcg tat cag cgc aca atc acc gat	48
Met Ser Ala Gln Pro Leu Pro Ala Ala Tyr Gln Arg Thr Ile Thr Asp	
1 5 10 15	
gtc att tcc atg cca aca ccg ggc cag gtt ccg ttt tct gta gag ttt	96
Val Ile Ser Met Pro Thr Pro Gly Gln Val Pro Phe Ser Val Glu Phe	
20 25 30	
atg ccg cca cga gat gag gca gca gaa gag cga ctc tgg aaa gcc gcc	144
Met Pro Pro Arg Asp Glu Ala Ala Glu Glu Arg Leu Trp Lys Ala Ala	
35 40 45	
gaa gca ttt cac gac tta gga gcc tct ttt gtc tcc gtt act tat ggt	192
Glu Ala Phe His Asp Leu Gly Ala Ser Phe Val Ser Val Thr Tyr Gly	
50 55 60	
gca ggc gga tct agc cgc gag cgc aca atg cgt gtc gcg cac aag ctt	240
Ala Gly Gly Ser Ser Arg Glu Arg Thr Met Arg Val Ala His Lys Leu	
65 70 75 80	
tct cgt cat ccg ttg acc acg ctc gtt cat ctc acg ctt gtg gaa cac	288
Ser Arg His Pro Leu Thr Thr Leu Val His Leu Thr Leu Val Glu His	
85 90 95	
acc caa gaa gaa tta gaa gaa att ctg tgc act tat gcg tcc cac ggg	336
Thr Gln Glu Glu Leu Glu Glu Ile Leu Cys Thr Tyr Ala Ser His Gly	
100 105 110	
ttg tct aac tta ctt gcc ttg cga ggc gat ccc cct ggc act gac ccg	384
Leu Ser Asn Leu Leu Ala Leu Arg Gly Asp Pro Pro Gly Thr Asp Pro	
115 120 125	
atg gct ccg tgg gtc cct acc gca ggc ggc cta gat tat gcc aaa gat	432
Met Ala Pro Trp Val Pro Thr Ala Gly Gly Leu Asp Tyr Ala Lys Asp	
130 135 140	
ttg atc gac ctc gtg cgc aag act gag cag acc tcg cac ttt cag gta	480
Leu Ile Asp Leu Val Arg Lys Thr Glu Gln Thr Ser His Phe Gln Val	
145 150 155 160	

M/43126

MetF

gga att gct agt ttc cca gaa ggg cac tac cga gcg cct agc att gag	528
Gly Ile Ala Ser Phe Pro Glu Gly His Tyr Arg Ala Pro Ser Ile Glu	
165 170 175	
gcg gat acg caa ttt aca ttg gaa aag ctg cga gct ggc gca gag ttt	576
Ala Asp Thr Gln Phe Thr Leu Glu Lys Leu Arg Ala Gly Ala Glu Phe	
180 185 190	
tcg att acc cag atg ttt ttt gat gtc gat cac tat tta cga ctg cga	624
Ser Ile Thr Gln Met Phe Phe Asp Val Asp His Tyr Leu Arg Leu Arg	
195 200 205	
gat cgc ttg gtt aag gcg gat cct gaa cat gga tca aag ccg atc atc	672
Asp Arg Leu Val Lys Ala Asp Pro Glu His Gly Ser Lys Pro Ile Ile	
210 215 220	
cca gga ctt atg ccc att acc agc ttg agg tcg gtt cgt agg cag atg	720
Pro Gly Leu Met Pro Ile Thr Ser Leu Arg Ser Val Arg Arg Gln Met	
225 230 235 240	
aa tta gca ggt gcc acc ttg cct aag gct tta gaa aaa cgg ctt ctc	768
lu Leu Ala Gly Ala Thr Leu Pro Lys Ala Leu Glu Lys Arg Leu Leu	
245 250 255	
gac gca gcg cgc ggc gat gag gaa gct cat cgc ggc gat att cgc aaa	816
Asp Ala Ala Arg Gly Asp Glu Glu Ala His Arg Gly Asp Ile Arg Lys	
260 265 270	
gta gga atc gaa gtc act act gag atg gca cag cgt ctt att tot gaa	864
Val Gly Ile Glu Val Thr Thr Glu Met Ala Gln Arg Leu Ile Ser Glu	
275 280 285	
ggg atc cca gac atc cat ttc atg acc atg aat tat gtt cga gcg acc	912
Gly Ile Pro Asp Ile His Phe Met Thr Met Asn Tyr Val Arg Ala Thr	
290 295 300	
caa gaa gta ctc cat aat ctc ggc atg gcg ccc gcg tgg gga aca cag	960
Gln Glu Val Leu His Asn Leu Gly Met Ala Pro Ala Trp Gly Thr Gln	
305 310 315 320	
aa ggc cac gac gct att cgc taa	984
ln Gly His Asp Ala Ile Arg	
325	

<210> 2

<211> 327

<212> PRT

<213> corynebacterium diphtheriae

<400> 2

Met Ser Ala Gln Pro Leu Pro Ala Ala Tyr Gln Arg Thr Ile Thr Asp	
1 5 10 15	
Val Ile Ser Met Pro Thr Pro Gly Gln Val Pro Phe Ser Val Glu Phe	
20 25 30	
Met Pro Pro Arg Asp Glu Ala Ala Glu Glu Arg Leu Trp Lys Ala Ala	
35 40 45	
Glu Ala Phe His Asp Leu Gly Ala Ser Phe Val Ser Val Thr Tyr Gly	
50 55 60	

Ala Gly Gly Ser Ser Arg Glu Arg Thr Met Arg Val Ala His Lys Leu
 65 70 75 80
 Ser Arg His Pro Leu Thr Thr Leu Val His Leu Thr Leu Val Glu His
 85 90 95
 Thr Gln Glu Glu Leu Glu Glu Ile Leu Cys Thr Tyr Ala Ser His Gly
 100 105 110
 Leu Ser Asn Leu Leu Ala Leu Arg Gly Asp Pro Pro Gly Thr Asp Pro
 115 120 125
 Met Ala Pro Trp Val Pro Thr Ala Gly Gly Leu Asp Tyr Ala Lys Asp
 130 135 140
 Leu Ile Asp Leu Val Arg Lys Thr Glu Gln Thr Ser His Phe Gln Val
 145 150 155 160
 Gly Ile Ala Ser Phe Pro Glu Gly His Tyr Arg Ala Pro Ser Ile Glu
 165 170 175
 Ala Asp Thr Gln Phe Thr Leu Glu Lys Leu Arg Ala Gly Ala Glu Phe
 180 185 190
 Ser Ile Thr Gln Met Phe Phe Asp Val Asp His Tyr Leu Arg Leu Arg
 195 200 205
 Asp Arg Leu Val Lys Ala Asp Pro Glu His Gly Ser Lys Pro Ile Ile
 210 215 220
 Pro Gly Leu Met Pro Ile Thr Ser Leu Arg Ser Val Arg Arg Gln Met
 225 230 235 240
 Glu Leu Ala Gly Ala Thr Leu Pro Lys Ala Leu Glu Lys Arg Leu Leu
 245 250 255
 Asp Ala Ala Arg Gly Asp Glu Glu Ala His Arg Gly Asp Ile Arg Lys
 260 265 270
 Val Gly Ile Glu Val Thr Thr Glu Met Ala Gln Arg Leu Ile Ser Glu
 275 280 285
 Gly Ile Pro Asp Ile His Phe Met Thr Met Asn Tyr Val Arg Ala Thr
 290 295 300
 Gln Glu Val Leu His Asn Leu Gly Met Ala Pro Ala Trp Gly Thr Gln
 305 310 315 320
 Gln Gly His Asp Ala Ile Arg
 325

<210> 3

<211> 924

<212> DNA

<213> Streptomyces lividans

<220>

<221> CDS

<222> (1) .. (921)

<223> RSV00084

<400> 3

atg gcc ctc gga acc gca agc acg agg acg gat cgc gcc cgc acg gtg	48
Met Ala Leu Gly Thr Ala Ser Thr Arg Thr Asp Arg Ala Arg Thr Val	
1 5 10 15	
cgt gac atc ctc gcc acc ggc aag acg acg tac tcg ttc gag ttc tcg	96
Arg Asp Ile Leu Ala Thr Gly Lys Thr Thr Tyr Ser Phe Glu Phe Ser	
20 25 30	
gcg ccg aag acg ccc aag ggc gag aag aac ctc tgg agc gcg ctg cgg	144
Ala Pro Lys Thr Pro Lys Gly Glu Lys Asn Leu Trp Ser Ala Leu Arg	
35 40 45	
cgg gtc gag gcc gtg gcc ccg gac ttc gtc tcc gtg acc tac ggc gcc	192
Arg Val Glu Ala Val Ala Pro Asp Phe Val Ser Val Thr Tyr Gly Ala	
50 55 60	
ggc ggc tcc acg cgc gcc ggc acg gtc cgc gag acc cag cag atc gtc	240
Gly Gly Ser Thr Arg Ala Gly Thr Val Arg Glu Thr Gln Gln Ile Val	
65 70 75 80	
cc gac acc acg ctg acc ccg gtg gcc cac ctc acc gcc gtc gac cac	288
Ala Asp Thr Thr Leu Thr Pro Val Ala His Leu Thr Ala Val Asp His	
85 90 95	
tcc gtc gcc gag ctg cgc aac atc atc ggc cag tac gcc gac gcc ggg	336
Ser Val Ala Glu Leu Arg Asn Ile Ile Gly Gln Tyr Ala Asp Ala Gly	
100 105 110	
atc cgc aac atg ctg gcc gtg cgc ggc gac ccg ccc ggc gac ccg aac	384
Ile Arg Asn Met Leu Ala Val Arg Gly Asp Pro Pro Gly Asp Pro Asn	
115 120 125	
gcc gac tgg atc gcg cac ccc gag ggc ctg acc tac gcg gcc gaa ctg	432
Ala Asp Trp Ile Ala His Pro Glu Gly Leu Thr Tyr Ala Ala Glu Leu	
130 135 140	
gtc agg ctc atc aag gag tcg gga gac ttc tgc gtc ggc gtc gcc gcc	480
Val Arg Leu Ile Lys Glu Ser Gly Asp Phe Cys Val Gly Val Ala Ala	
145 150 155 160	
tc ccc gag atg cac ccg cgc tcc gcc gac tgg gac acg gac gtc acg	528
Phe Pro Glu Met His Pro Arg Ser Ala Asp Trp Asp Thr Asp Val Thr	
165 170 175	
aac ttc gtc gac aag tgc cgg gcc ggc gcc gac tac gcc atc acc cag	576
Asn Phe Val Asp Lys Cys Arg Ala Gly Ala Asp Tyr Ala Ile Thr Gln	
180 185 190	
atg ttc ttc cag ccc gac tcc tac ctc cgg ctg cgc gac cgg gtc gcc	624
Met Phe Phe Gln Pro Asp Ser Tyr Leu Arg Leu Arg Asp Arg Val Ala	
195 200 205	
gcg gcc ggc tgc gcg acc ccg gtc att ccc gag gtc atg ccg gtg acc	672
Ala Ala Gly Cys Ala Thr Pro Val Ile Pro Glu Val Met Pro Val Thr	
210 215 220	
agt gtg aag atg ctg gag agg ttg ccg aag ctc agc aac gcc tcg ttc	720
Ser Val Lys Met Leu Glu Arg Leu Pro Lys Leu Ser Asn Ala Ser Phe	
225 230 235 240	
ccg gcg gag ctg aaa gag cgg atc ctc aca gcc aag gac gat ccg gcg	768
Pro Ala Glu Leu Lys Glu Arg Ile Leu Thr Ala Lys Asp Asp Pro Ala	

245										250					255					
gct gta cgc tcg atc ggc atc gag ttc gcc acg gag ttc tgc gcg cgg	816																			
Ala Val Arg Ser Ile Gly Ile Glu Phe Ala Thr Glu Phe Cys Ala Arg																				
260 265 270																				
ctg ctg gcc gag gga gtg cca gga ctg cac ttc atc acg ctc aac aac	864																			
Leu Leu Ala Glu Gly Val Pro Gly Leu His Phe Ile Thr Leu Asn Asn																				
275 280 285																				
tcc acg gcg acg ctg gaa atc tac gag aac ctg ggc ctg cac cac cca	912																			
Ser Thr Ala Thr Leu Glu Ile Tyr Glu Asn Leu Gly Leu His His Pro																				
290 295 300																				
ccg cgg gcc tag	924																			
Pro Arg Ala																				
305																				

<210> 4

<211> 307

<212> PRT

<213> Streptomyces lividans

<400> 4

Met Ala Leu Gly Thr Ala Ser Thr Arg Thr Asp Arg Ala Arg Thr Val	
1 5 10 15	
Arg Asp Ile Leu Ala Thr Gly Lys Thr Thr Tyr Ser Phe Glu Phe Ser	
20 25 30	
Ala Pro Lys Thr Pro Lys Gly Glu Lys Asn Leu Trp Ser Ala Leu Arg	
35 40 45	
Arg Val Glu Ala Val Ala Pro Asp Phe Val Ser Val Thr Tyr Gly Ala	
50 55 60	
Gly Gly Ser Thr Arg Ala Gly Thr Val Arg Glu Thr Gln Gln Ile Val	
65 70 75 80	
Ala Asp Thr Thr Leu Thr Pro Val Ala His Leu Thr Ala Val Asp His	
85 90 95	
Ser Val Ala Glu Leu Arg Asn Ile Ile Gly Gln Tyr Ala Asp Ala Gly	
100 105 110	
Ile Arg Asn Met Leu Ala Val Arg Gly Asp Pro Pro Gly Asp Pro Asn	
115 120 125	
Ala Asp Trp Ile Ala His Pro Glu Gly Leu Thr Tyr Ala Ala Glu Leu	
130 135 140	
Val Arg Leu Ile Lys Glu Ser Gly Asp Phe Cys Val Gly Val Ala Ala	
145 150 155 160	
Phe Pro Glu Met His Pro Arg Ser Ala Asp Trp Asp Thr Asp Val Thr	
165 170 175	
Asn Phe Val Asp Lys Cys Arg Ala Gly Ala Asp Tyr Ala Ile Thr Gln	
180 185 190	
Met Phe Phe Gln Pro Asp Ser Tyr Leu Arg Leu Arg Asp Arg Val Ala	
195 200 205	

Ala Ala Gly Cys Ala Thr Pro Val Ile Pro Glu Val Met Pro Val Thr
 210 215 220

Ser Val Lys Met Leu Glu Arg Leu Pro Lys Leu Ser Asn Ala Ser Phe
 225 230 235 240

Pro Ala Glu Leu Lys Glu Arg Ile Leu Thr Ala Lys Asp Asp Pro Ala
 245 250 255

Ala Val Arg Ser Ile Gly Ile Glu Phe Ala Thr Glu Phe Cys Ala Arg
 260 265 270

Leu Leu Ala Glu Gly Val Pro Gly Leu His Phe Ile Thr Leu Asn Asn
 275 280 285

Ser Thr Ala Thr Leu Glu Ile Tyr Glu Asn Leu Gly Leu His His Pro
 290 295 300

Pro Arg Ala
 305

<210> 5
 <211> 924
 <212> DNA
 <213> Streptomyces coelicolor

<220>
 <221> CDS
 <222> (1)..(921)
 <223> RSX01699

<400> 5
 atg gcc ctc gga acc gca agc acg agg acg gat cgc gcc cgc acg gtg 48
 Met Ala Leu Gly Thr Ala Ser Thr Arg Thr Asp Arg Ala Arg Thr Val
 1 5 10 15

cgt gac atc ctc gcc acc ggc aag acg acg tac tcg ttc gag ttc tcg 96
 Arg Asp Ile Leu Ala Thr Gly Lys Thr Thr Tyr Ser Phe Glu Phe Ser
 20 25 30

cgc ccg aag acg ccc aag ggc gag agg aac ctc tgg agc gcg ctg cgg 144
 Ala Pro Lys Thr Pro Lys Gly Glu Arg Asn Leu Trp Ser Ala Leu Arg
 35 40 45

cgg gtc gag gcc gtg gcc ccg gac ttc gtc tcc gtg acc tac ggc gcc 192
 Arg Val Glu Ala Val Ala Pro Asp Phe Val Ser Val Thr Tyr Gly Ala
 50 55 60

ggc ggc tcc acg cgc gcc ggc acg gtc cgc gag acc cag cag atc gtc 240
 Gly Gly Ser Thr Arg Ala Gly Thr Val Arg Glu Thr Gln Gln Ile Val
 65 70 75 80

gcc gac acc acg ctg acc ccg gtg gcc cac ctc acc gcc gtc gac cac 288
 Ala Asp Thr Thr Leu Thr Pro Val Ala His Leu Thr Ala Val Asp His
 85 90 95

tcc gtc gcc gag ctg cgc aac atc atc ggc cag tac gcc gac gcc ggg 336
 Ser Val Ala Glu Leu Arg Asn Ile Ile Gly Gln Tyr Ala Asp Ala Gly
 100 105 110

atc cgc aac atg ctg gcc gtg cgc ggc gac ccg ccc ggc gac ccg aac 384

```

Ile Arg Asn Met Leu Ala Val Arg Gly Asp Pro Pro Gly Asp Pro Asn
115                               120                               125

gcc gac tgg atc gcg cac ccc gag ggc ctg acc tac gcg gcc gaa ctg 432
Ala Asp Trp Ile Ala His Pro Glu Gly Leu Thr Tyr Ala Ala Glu Leu
130                               135                               140

gtc agg ctc atc aag gag tgc ggc gac ttc tgc gtc ggc gtc gcg gcc 480
Val Arg Leu Ile Lys Glu Ser Gly Asp Phe Cys Val Gly Val Ala Ala
145                               150                               155                               160

ttc ccc gag atg cac ccg cgc tcc gcc gac tgg gac acg gac gtc acg 528
Phe Pro Glu Met His Pro Arg Ser Ala Asp Trp Asp Thr Asp Val Thr
165                               170                               175

aac ttc gtc gac aag tgc cgg gcc ggc gcc gac tac gcc atc acc cag 576
Asn Phe Val Asp Lys Cys Arg Ala Gly Ala Asp Tyr Ala Ile Thr Gln
180                               185                               190

atg ttc ttc cag ccc gac tcc tat ctc cgg ctg cgc gac cgg gtc gcc 624
Met Phe Phe Gln Pro Asp Ser Tyr Leu Arg Leu Arg Asp Arg Val Ala
195                               200                               205

gcg gcc ggc tgc gcg acc ccg gtc atc ccc gag gtc atg ccg gtg acc 672
Ala Ala Gly Cys Ala Thr Pro Val Ile Pro Glu Val Met Pro Val Thr
210                               215                               220

agt gtg aag atg ctg gag agg ttg ccg aag ctc agc aac gcc tcg ttc 720
Ser Val Lys Met Leu Glu Arg Leu Pro Lys Leu Ser Asn Ala Ser Phe
225                               230                               235                               240

ccg gcg gag ttg aaa gag cgg atc ctc aca gcc aag gac gat ccg gcg 768
Pro Ala Glu Leu Lys Glu Arg Ile Leu Thr Ala Lys Asp Asp Pro Ala
245                               250                               255

gct gta cgc tgc atc ggc atc gag ttc gcc acg gag ttc tgc gcg cgg 816
Ala Val Arg Ser Ile Gly Ile Glu Phe Ala Thr Glu Phe Cys Ala Arg
260                               265                               270

ctg ctg gcc gag gga gtg cca gga ctg cac ttc atc acg ctc aac aac 864
Leu Leu Ala Glu Gly Val Pro Gly Leu His Phe Ile Thr Leu Asn Asn
275                               280                               285

tcc acg gcg acg ctg gaa atc tac gag aac ctg ggc ctg cac cac cca 912
Ser Thr Ala Thr Leu Glu Ile Tyr Glu Asn Leu Gly Leu His His Pro
290                               295                               300

ccg cgg gcc tag 924
Pro Arg Ala
305

```

<210> 6

<211> 307

<212> PRT

<213> Streptomyces coelicolor

<400> 6

```

Met Ala Leu Gly Thr Ala Ser Thr Arg Thr Asp Arg Ala Arg Thr Val
  1              5              10              15

```

```

Arg Asp Ile Leu Ala Thr Gly Lys Thr Thr Tyr Ser Phe Glu Phe Ser

```

M/43126

MetF

20					25					30					
Ala	Pro	Lys	Thr	Pro	Lys	Gly	Glu	Arg	Asn	Leu	Trp	Ser	Ala	Leu	Arg
		35					40					45			
Arg	Val	Glu	Ala	Val	Ala	Pro	Asp	Phe	Val	Ser	Val	Thr	Tyr	Gly	Ala
	50					55					60				
Gly	Gly	Ser	Thr	Arg	Ala	Gly	Thr	Val	Arg	Glu	Thr	Gln	Gln	Ile	Val
65					70					75				80	
Ala	Asp	Thr	Thr	Leu	Thr	Pro	Val	Ala	His	Leu	Thr	Ala	Val	Asp	His
				85					90					95	
Ser	Val	Ala	Glu	Leu	Arg	Asn	Ile	Ile	Gly	Gln	Tyr	Ala	Asp	Ala	Gly
			100					105					110		
Ile	Arg	Asn	Met	Leu	Ala	Val	Arg	Gly	Asp	Pro	Pro	Gly	Asp	Pro	Asn
		115					120					125			
Ala	Asp	Trp	Ile	Ala	His	Pro	Glu	Gly	Leu	Thr	Tyr	Ala	Ala	Glu	Leu
	130					135					140				
Val	Arg	Leu	Ile	Lys	Glu	Ser	Gly	Asp	Phe	Cys	Val	Gly	Val	Ala	Ala
145				150						155					160
Phe	Pro	Glu	Met	His	Pro	Arg	Ser	Ala	Asp	Trp	Asp	Thr	Asp	Val	Thr
				165					170					175	
Asn	Phe	Val	Asp	Lys	Cys	Arg	Ala	Gly	Ala	Asp	Tyr	Ala	Ile	Thr	Gln
			180					185					190		
Met	Phe	Phe	Gln	Pro	Asp	Ser	Tyr	Leu	Arg	Leu	Arg	Asp	Arg	Val	Ala
		195					200					205			
Ala	Ala	Gly	Cys	Ala	Thr	Pro	Val	Ile	Pro	Glu	Val	Met	Pro	Val	Thr
		210				215					220				
Ser	Val	Lys	Met	Leu	Glu	Arg	Leu	Pro	Lys	Leu	Ser	Asn	Ala	Ser	Phe
225				230						235				240	
Pro	Ala	Glu	Leu	Lys	Glu	Arg	Ile	Leu	Thr	Ala	Lys	Asp	Asp	Pro	Ala
			245					250						255	
Ala	Val	Arg	Ser	Ile	Gly	Ile	Glu	Phe	Ala	Thr	Glu	Phe	Cys	Ala	Arg
			260					265					270		
Leu	Leu	Ala	Glu	Gly	Val	Pro	Gly	Leu	His	Phe	Ile	Thr	Leu	Asn	Asn
		275					280					285			
Ser	Thr	Ala	Thr	Leu	Glu	Ile	Tyr	Glu	Asn	Leu	Gly	Leu	His	His	Pro
		290				295					300				
Pro	Arg	Ala													
305															

<210> 7

<211> 891

<212> DNA

<213> Aquifex aeolicus

<220>

M/43126

MetF

<221> CDS

<222> (1) .. (888)

<223> RAA00346

<400> 7

atg	aaa	ata	gga	gat	ata	ctg	agg	aaa	gga	gtt	ttc	agt	att	tct	ttt	48
Met	Lys	Ile	Gly	Asp	Ile	Leu	Arg	Lys	Gly	Val	Phe	Ser	Ile	Ser	Phe	
1				5					10					15		
gag	ttc	ttt	cca	ccg	aag	act	gaa	gag	gga	gaa	aga	cag	ctc	ttt	gaa	96
Glu	Phe	Phe	Pro	Pro	Lys	Thr	Glu	Glu	Gly	Glu	Arg	Gln	Leu	Phe	Glu	
			20					25					30			
act	ata	agg	aaa	ctt	gag	aaa	tta	aat	cct	act	ttt	gta	tcc	gtt	act	144
Thr	Ile	Arg	Lys	Leu	Glu	Lys	Leu	Asn	Pro	Thr	Phe	Val	Ser	Val	Thr	
		35					40					45				
tac	ggg	gca	ggg	ggg	tcg	act	aga	gat	aga	act	agg	aat	ata	gta	cag	192
Tyr	Gly	Ala	Gly	Gly	Ser	Thr	Arg	Asp	Arg	Thr	Arg	Asn	Ile	Val	Gln	
	50					55					60					
ata	ata	cac	gag	gaa	act	aac	ctc	acc	gtt	atg	gca	cac	ctc	acc	tgt	240
Lys	Ile	His	Glu	Glu	Thr	Asn	Leu	Thr	Val	Met	Ala	His	Leu	Thr	Cys	
65					70				75					80		
ata	gca	cac	acg	aga	gag	gag	ctt	att	gat	atc	ctt	caa	gat	tac	aaa	288
Ile	Ala	His	Thr	Arg	Glu	Glu	Leu	Ile	Asp	Ile	Leu	Gln	Asp	Tyr	Lys	
			85					90						95		
aac	ata	ggg	ata	gag	aac	att	ctc	gct	ttg	agg	ggg	gac	gtt	ccg	agg	336
Asn	Ile	Gly	Ile	Glu	Asn	Ile	Leu	Ala	Leu	Arg	Gly	Asp	Val	Pro	Arg	
			100					105					110			
gac	aaa	ccg	gac	tgg	aga	ccg	ccg	aag	ggg	gcg	tgc	aag	tat	gca	aaa	384
Asp	Lys	Pro	Asp	Trp	Arg	Pro	Pro	Lys	Gly	Ala	Cys	Lys	Tyr	Ala	Lys	
		115					120					125				
gag	ctc	gta	gaa	ctg	atc	agg	aag	gag	ttc	gga	gac	tgg	ttt	tct	atc	432
Glu	Leu	Val	Glu	Leu	Ile	Arg	Lys	Glu	Phe	Gly	Asp	Trp	Phe	Ser	Ile	
	130					135					140					
gaa	gtg	gct	tct	tat	cct	gaa	gga	cat	ccg	gaa	tca	ccg	aac	ctc	gag	480
Gly	Val	Ala	Ser	Tyr	Pro	Glu	Gly	His	Pro	Glu	Ser	Pro	Asn	Leu	Glu	
145					150				155					160		
tgg	gaa	gtg	aag	tac	ttt	aag	gaa	aag	gta	gag	gca	ggg	gca	gac	ttc	528
Trp	Glu	Val	Lys	Tyr	Phe	Lys	Glu	Lys	Val	Glu	Ala	Gly	Ala	Asp	Phe	
			165					170						175		
tcg	att	act	cag	atg	ttt	ttc	gtg	aac	gat	tac	tac	tac	agg	ttt	gtg	576
Ser	Ile	Thr	Gln	Met	Phe	Phe	Val	Asn	Asp	Tyr	Tyr	Tyr	Arg	Phe	Val	
			180					185					190			
gaa	atg	tgc	aaa	aat	gca	ggg	ata	gat	ata	tct	ata	att	ccg	gga	att	624
Glu	Met	Cys	Lys	Asn	Ala	Gly	Ile	Asp	Ile	Ser	Ile	Ile	Pro	Gly	Ile	
		195				200						205				
atg	cct	att	act	aac	ttc	aaa	cag	ata	aga	aag	ttt	gct	tct	ctt	tgc	672
Met	Pro	Ile	Thr	Asn	Phe	Lys	Gln	Ile	Arg	Lys	Phe	Ala	Ser	Leu	Cys	
	210					215					220					
gga	gag	acg	att	cca	cag	agt	ctt	ata	gaa	aag	ctt	gaa	aaa	gtg	gag	720
Gly	Ala	Thr	Ile	Pro	Gln	Ser	Leu	Ile	Glu	Lys	Leu	Glu	Lys	Val	Glu	

225	230	235	240	
gat aaa ccg gaa gaa gta aaa aag ata ggg att gag ttt gcc ata aat				768
Asp Lys Pro Glu Glu Val Lys Lys Ile Gly Ile Glu Phe Ala Ile Asn	245	250	255	
cag tgt ttg gat ctc ata gaa cac gga gtt ccg ggg ctt cac ttc tac				816
Gln Cys Leu Asp Leu Ile Glu His Gly Val Pro Gly Leu His Phe Tyr	260	265	270	
act ctg aac aag tcc gac gca act ttg aag ata tac gag gct ata aag				864
Thr Leu Asn Lys Ser Asp Ala Thr Leu Lys Ile Tyr Glu Ala Ile Lys	275	280	285	
gat aaa ata ccg gcc cgt tca act taa				891
Asp Lys Ile Pro Ala Arg Ser Thr	290	295		

<210> 8

<211> 296

<212> PRT

<213> Aquifex aeolicus

<400> 8

Met Lys Ile Gly Asp Ile Leu Arg Lys Gly Val Phe Ser Ile Ser Phe	1	5	10	15
Glu Phe Phe Pro Pro Lys Thr Glu Glu Gly Glu Arg Gln Leu Phe Glu	20	25	30	
Thr Ile Arg Lys Leu Glu Lys Leu Asn Pro Thr Phe Val Ser Val Thr	35	40	45	
Tyr Gly Ala Gly Gly Ser Thr Arg Asp Arg Thr Arg Asn Ile Val Gln	50	55	60	
Lys Ile His Glu Glu Thr Asn Leu Thr Val Met Ala His Leu Thr Cys	65	70	75	80
Ile Ala His Thr Arg Glu Glu Leu Ile Asp Ile Leu Gln Asp Tyr Lys	85	90	95	
Asn Ile Gly Ile Glu Asn Ile Leu Ala Leu Arg Gly Asp Val Pro Arg	100	105	110	
Asp Lys Pro Asp Trp Arg Pro Pro Lys Gly Ala Cys Lys Tyr Ala Lys	115	120	125	
Glu Leu Val Glu Leu Ile Arg Lys Glu Phe Gly Asp Trp Phe Ser Ile	130	135	140	
Gly Val Ala Ser Tyr Pro Glu Gly His Pro Glu Ser Pro Asn Leu Glu	145	150	155	160
Trp Glu Val Lys Tyr Phe Lys Glu Lys Val Glu Ala Gly Ala Asp Phe	165	170	175	
Ser Ile Thr Gln Met Phe Phe Val Asn Asp Tyr Tyr Tyr Arg Phe Val	180	185	190	
Glu Met Cys Lys Asn Ala Gly Ile Asp Ile Ser Ile Ile Pro Gly Ile	195	200	205	

Met Pro Ile Thr Asn Phe Lys Gln Ile Arg Lys Phe Ala Ser Leu Cys
 210 215 220

Gly Ala Thr Ile Pro Gln Ser Leu Ile Glu Lys Leu Glu Lys Val Glu
 225 230 235 240

Asp Lys Pro Glu Glu Val Lys Lys Ile Gly Ile Glu Phe Ala Ile Asn
 245 250 255

Gln Cys Leu Asp Leu Ile Glu His Gly Val Pro Gly Leu His Phe Tyr
 260 265 270

Thr Leu Asn Lys Ser Asp Ala Thr Leu Lys Ile Tyr Glu Ala Ile Lys
 275 280 285

Asp Lys Ile Pro Ala Arg Ser Thr
 290 295

<210> 9
 <211> 831
 <212> DNA
 <213> Burkholderia cepacia

<220>
 <221> CDS
 <222> (1)..(828)
 <223> RBU14992

<400> 9
 atg aac ccg atc gaa ctt tca ttc gaa ttc ttc ccg ccg aaa acg cag 48

Met Asn Pro Ile Glu Leu Ser Phe Glu Phe Phe Pro Pro Lys Thr Gln
 1 5 10 15

gaa ggc gtg gac aag ctg cgc gcc acg cgc gcc cag ctc gcc acg ctc 96
 Glu Gly Val Asp Lys Leu Arg Ala Thr Arg Ala Gln Leu Ala Thr Leu
 20 25 30

aag ccc aag ttc gtg tcc gtc acg ttc ggc gcc ggc ggc tcg acg caa 144
 Lys Pro Lys Phe Val Ser Val Thr Phe Gly Ala Gly Gly Ser Thr Gln
 35 40 45

cag ggc acg ctc gac acc gtc gtc gat atg gcg aag gaa ggg ctc gaa 192
 Gln Gly Thr Leu Asp Thr Val Val Asp Met Ala Lys Glu Gly Leu Glu
 50 55 60

gcg gcg ccg cac gtg tcg tgc atc ggc tcg tcg aaa gag agc ctg cgc 240
 Ala Ala Pro His Val Ser Cys Ile Gly Ser Ser Lys Glu Ser Leu Arg
 65 70 75 80

gcc att ctc aac gag tac cgc gca cat ggc atc cgc cat atc gtc gcg 288

Ala Ile Leu Asn Glu Tyr Arg Ala His Gly Ile Arg His Ile Val Ala
 85 90 95

ctg cgc ggc gat ctg ccg tcc ggc atg ggc gaa gtc ggc gag ctg cgc 336
 Leu Arg Gly Asp Leu Pro Ser Gly Met Gly Glu Val Gly Glu Leu Arg
 100 105 110

tat gcg tcg gaa ctg gtg agc ttt atc cgc gcc gaa ttc ggc gac tgg 384
 Tyr Ala Ser Glu Leu Val Ser Phe Ile Arg Ala Glu Phe Gly Asp Trp

115	120	125	
ttc tgc atc gag gtg gcc ggc tat ccg gaa tac cac ccg cag tcg cgc			432
Phe Cys Ile Glu Val Ala Gly Tyr Pro Glu Tyr His Pro Gln Ser Arg			
130	135	140	
tcg ccg cgt cag gat ctg gaa aac ttc gcc cgc aag gtg aag gcc ggc			480
Ser Pro Arg Gln Asp Leu Glu Asn Phe Ala Arg Lys Val Lys Ala Gly			
145	150	155	160
gcc aat tcg gcg atc aca cag tac ttc ttc aat gca gac gcg tat ttc			528
Ala Asn Ser Ala Ile Thr Gln Tyr Phe Phe Asn Ala Asp Ala Tyr Phe			
	165	170	175
cgt ttc gtc gac gac gcg aga aag ctc ggc gtg gac gtg ccg atc gtg			576
Arg Phe Val Asp Asp Ala Arg Lys Leu Gly Val Asp Val Pro Ile Val			
	180	185	190
ccg ggc atc atg ccg atc acg aac ttc tcg cag ctg atg cgt ttc tcg			624
Pro Gly Ile Met Pro Ile Thr Asn Phe Ser Gln Leu Met Arg Phe Ser			
	195	200	205
agc atg tgc ggc gct gaa gtg cca cgc tgg atc gcg cgc cgg ctg gaa			672
Glu Met Cys Gly Ala Glu Val Pro Arg Trp Ile Ala Arg Arg Leu Glu			
	210	215	220
agc ttc ggc gac gat cgc gag tca att cgc gcg ttc ggg ctg gat gtg			720
Ser Phe Gly Asp Asp Arg Glu Ser Ile Arg Ala Phe Gly Leu Asp Val			
225	230	235	240
gtg acg gac ctg tgc agg cgt ctg atc gat gcg aag gtg ccg ggc ctg			768
Val Thr Asp Leu Cys Arg Arg Leu Ile Asp Ala Lys Val Pro Gly Leu			
	245	250	255
cac ttc tac acg cta aac ggc gca gcg gcg acc aag gcg atc tgc gaa			816
His Phe Tyr Thr Leu Asn Gly Ala Ala Ala Thr Lys Ala Ile Cys Glu			
	260	265	270
cgg ttg aac gtt taa			831
Arg Leu Asn Val			
275			

<210> 10

<211> 276

<212> PRT

<213> Burkholderia cepacia

<400> 10

Met	Asn	Pro	Ile	Glu	Leu	Ser	Phe	Glu	Phe	Phe	Pro	Pro	Lys	Thr	Gln
1				5				10						15	

Glu	Gly	Val	Asp	Lys	Leu	Arg	Ala	Thr	Arg	Ala	Gln	Leu	Ala	Thr	Leu
		20						25				30			

Lys	Pro	Lys	Phe	Val	Ser	Val	Thr	Phe	Gly	Ala	Gly	Gly	Ser	Thr	Gln
		35					40				45				

Gln	Gly	Thr	Leu	Asp	Thr	Val	Val	Asp	Met	Ala	Lys	Glu	Gly	Leu	Glu
	50					55					60				

Ala Ala Pro His Val Ser Cys Ile Gly Ser Ser Lys Glu Ser Leu Arg

M/43126

MetF

65	70	75	80
Ala Ile Leu Asn Glu Tyr Arg Ala His Gly Ile Arg His Ile Val Ala	85	90	95
Leu Arg Gly Asp Leu Pro Ser Gly Met Gly Glu Val Gly Glu Leu Arg	100	105	110
Tyr Ala Ser Glu Leu Val Ser Phe Ile Arg Ala Glu Phe Gly Asp Trp	115	120	125
Phe Cys Ile Glu Val Ala Gly Tyr Pro Glu Tyr His Pro Gln Ser Arg	130	135	140
Ser Pro Arg Gln Asp Leu Glu Asn Phe Ala Arg Lys Val Lys Ala Gly	145	150	155
Ala Asn Ser Ala Ile Thr Gln Tyr Phe Phe Asn Ala Asp Ala Tyr Phe	165	170	175
Arg Phe Val Asp Asp Ala Arg Lys Leu Gly Val Asp Val Pro Ile Val	180	185	190
Pro Gly Ile Met Pro Ile Thr Asn Phe Ser Gln Leu Met Arg Phe Ser	195	200	205
Glu Met Cys Gly Ala Glu Val Pro Arg Trp Ile Ala Arg Arg Leu Glu	210	215	220
Ser Phe Gly Asp Asp Arg Glu Ser Ile Arg Ala Phe Gly Leu Asp Val	225	230	235
Val Thr Asp Leu Cys Arg Arg Leu Ile Asp Ala Lys Val Pro Gly Leu	245	250	255
His Phe Tyr Thr Leu Asn Gly Ala Ala Ala Thr Lys Ala Ile Cys Glu	260	265	270
Arg Leu Asn Val	275		

<210> 11
 <211> 846
 <212> DNA
 <213> Nitrosomonas europaea

<220>
 <221> CDS
 <222> (1)..(843)
 <223> RNE02657

<400> 11

atg caa tcc cag aaa aaa ttt acc ccc aca ttc agt ttt gaa ttt ttc	48
Met Gln Ser Gln Lys Lys Phe Thr Pro Thr Phe Ser Phe Glu Phe Phe	
1 5 10 15	
ccg ccg cag aca ccg gaa ggc atg gaa aag ctg cgg gca acg cgc ata	96
Pro Pro Gln Thr Pro Glu Gly Met Glu Lys Leu Arg Ala Thr Arg Ile	
20 25 30	
cag ctt gct cag ttc aat ccg aag ttt ttt tcg gtg acg ttt ggt gcc	144

Gln	Leu	Ala	Gln	Phe	Asn	Pro	Lys	Phe	Phe	Ser	Val	Thr	Phe	Gly	Ala	
		35					40					45				
ggc	gga	tcc	act	cgt	gaa	cgc	acg	ctc	gaa	acc	gtg	ctg	gaa	att	cag	192
Gly	Gly	Ser	Thr	Arg	Glu	Arg	Thr	Leu	Glu	Thr	Val	Leu	Glu	Ile	Gln	
	50					55					60					
gca	gaa	ggc	tat	ccg	gta	gcg	ccc	cat	ctt	tcc	tgt	atc	ggc	tcc	acg	240
Ala	Glu	Gly	Tyr	Pro	Val	Ala	Pro	His	Leu	Ser	Cys	Ile	Gly	Ser	Thr	
	65				70					75					80	
cgt	gac	aat	atc	cgt	tcg	atc	ctt	gag	aaa	tat	cac	agt	cac	ggc	atc	288
Arg	Asp	Asn	Ile	Arg	Ser	Ile	Leu	Glu	Lys	Tyr	His	Ser	His	Gly	Ile	
				85					90					95		
agc	cgc	att	gtg	gcg	cta	cgt	ggc	gat	tta	ccc	tcc	ggc	atg	gcg	cag	336
Ser	Arg	Ile	Val	Ala	Leu	Arg	Gly	Asp	Leu	Pro	Ser	Gly	Met	Ala	Gln	
			100					105					110			
gcg	gga	gaa	ttc	cgc	tac	gcc	aac	gag	ctg	gta	gca	ttt	atc	cgc	aag	384
Ala	Gly	Glu	Phe	Arg	Tyr	Ala	Asn	Glu	Leu	Val	Ala	Phe	Ile	Arg	Lys	
		115					120					125				
gag	ttc	ggc	gat	acc	ttc	tgg	atc	gaa	gtg	gcg	gct	tat	ccg	gaa	tat	432
Glu	Phe	Gly	Asp	Thr	Phe	Trp	Ile	Glu	Val	Ala	Ala	Tyr	Pro	Glu	Tyr	
	130					135					140					
cat	cca	caa	gcc	cgc	tcc	gct	ctg	gag	gat	ttc	acc	aat	ttc	aga	cga	480
His	Pro	Gln	Ala	Arg	Ser	Ala	Leu	Glu	Asp	Phe	Thr	Asn	Phe	Arg	Arg	
	145				150					155					160	
aaa	gtc	gaa	gca	ggc	tcc	aat	gca	gcg	att	acc	cag	ttt	ttc	tat	aac	528
Lys	Val	Glu	Ala	Gly	Ser	Asn	Ala	Ala	Ile	Thr	Gln	Phe	Phe	Tyr	Asn	
				165					170					175		
gtg	gat	gcc	tat	ctg	cat	ttc	gta	gag	atg	tgt	gaa	gct	gcg	gat	ctg	576
Val	Asp	Ala	Tyr	Leu	His	Phe	Val	Glu	Met	Cys	Glu	Ala	Ala	Asp	Leu	
			180					185					190			
aat	atc	ccg	atc	gtt	ccc	ggc	atc	atg	ccg	atc	agc	aaa	ttt	tct	caa	624
Asn	Ile	Pro	Ile	Val	Pro	Gly	Ile	Met	Pro	Ile	Ser	Lys	Phe	Ser	Gln	
		195				200						205				
ctg	gca	aga	ttt	tcg	gat	ggc	tgt	gga	gca	gaa	att	cca	cgc	tgg	att	672
Leu	Ala	Arg	Phe	Ser	Asp	Gly	Cys	Gly	Ala	Glu	Ile	Pro	Arg	Trp	Ile	
	210					215					220					
cgc	aga	aaa	ctg	gaa	agc	ttc	ggc	gat	gat	att	ccg	tct	atc	cag	gca	720
Arg	Arg	Lys	Leu	Glu	Ser	Phe	Gly	Asp	Asp	Ile	Pro	Ser	Ile	Gln	Ala	
	225				230					235					240	
ttc	ggg	ctg	gat	gtc	gtc	aca	gcg	tta	tgt	gct	cgt	ctg	ctg	gaa	gcc	768
Phe	Gly	Leu	Asp	Val	Val	Thr	Ala	Leu	Cys	Ala	Arg	Leu	Leu	Glu	Ala	
				245					250					255		
ggc	gca	ccc	ggc	ctg	cat	ttc	tac	aca	ctc	aac	tcc	gcc	gta	cta	ccc	816
Gly	Ala	Pro	Gly	Leu	His	Phe	Tyr	Thr	Leu	Asn	Ser	Ala	Val	Leu	Pro	
			260					265					270			
aca	aaa	atc	tgg	caa	cgc	ctg	ggg	tta	tag							846
Thr	Lys	Ile	Trp	Gln	Arg	Leu	Gly	Leu								
		275					280									

```
<210> 12
<211> 281
<212> PRT
<213> Nitrosomonas europaea
```

<400> 12

Met Gln Ser Gln Lys Lys Phe Thr Pro Thr Phe Ser Phe Glu Phe Phe
1 5 10 15

Pro Pro Gln Thr Pro Glu Gly Met Glu Lys Leu Arg Ala Thr Arg Ile
20 25 30

Gln Leu Ala Gln Phe Asn Pro Lys Phe Phe Ser Val Thr Phe Gly Ala
35 40 45

Gly Gly Ser Thr Arg Glu Arg Thr Leu Glu Thr Val Leu Glu Ile Gln
50 55 60

Ala Glu Gly Tyr Pro Val Ala Pro His Leu Ser Cys Ile Gly Ser Thr
65 70 75 80

Arg Asp Asn Ile Arg Ser Ile Leu Glu Lys Tyr His Ser His Gly Ile
85 90 95

Ser Arg Ile Val Ala Leu Arg Gly Asp Leu Pro Ser Gly Met Ala Gln
100 105 110

Ala Gly Glu Phe Arg Tyr Ala Asn Glu Leu Val Ala Phe Ile Arg Lys
115 120 125

Glu Phe Gly Asp Thr Phe Trp Ile Glu Val Ala Ala Tyr Pro Glu Tyr
130 135 140

His Pro Gln Ala Arg Ser Ala Leu Glu Asp Phe Thr Asn Phe Arg Arg
145 150 155 160

Lys Val Glu Ala Gly Ser Asn Ala Ala Ile Thr Gln Phe Phe Tyr Asn
165 170 175

Val Asp Ala Tyr Leu His Phe Val Glu Met Cys Glu Ala Ala Asp Leu
180 185 190

```

Asn Ile Pro Ile Val Pro Gly Ile Met Pro Ile Ser Lys Phe Ser Gln
      195                               200                     205

```

Leu Ala Arg Phe Ser Asp Gly Cys Gly Ala Glu Ile Pro Arg Trp Ile
210 215 220

Arg Arg Lys Leu Glu Ser Phe Gly Asp Asp Ile Pro Ser Ile Gln Ala
225 230 235 240

Phe Gly Leu Asp Val Val Thr Ala Leu Cys Ala Arg Leu Leu Glu Ala
245 250 255

Gly Ala Pro Gly Leu His Phe Tyr Thr Leu Asn Ser Ala Val Leu Pro
260 265 270

Thr Lys Ile Trp Gln Arg Leu Gly Leu
275 280

<210> 13
 <211> 873
 <212> DNA
 <213> *Pseudomonas aeruginosa*

<220>
 <221> CDS
 <222> (1)..(870)
 <223> RPA03308

<400> 13
 gtg gtc gcg tcc aag gaa ccg atc atg agt cag agc gaa cgc cgt ttc 48
 Val Val Ala Ser Lys Glu Pro Ile Met Ser Gln Ser Glu Arg Arg Phe
 1 5 10 15
 agc ttc gag ttc ttc ccg gcg aag acc gag gcc ggc cat gaa aag ctg 96
 Ser Phe Glu Phe Phe Pro Ala Lys Thr Glu Ala Gly His Glu Lys Leu
 20 25 30
 ttg gcc acc gcc cgc aac ctg gcg ggc tac aag ccc gac ttc ttc tcc 144
 Leu Ala Thr Ala Arg Asn Leu Ala Gly Tyr Lys Pro Asp Phe Phe Ser
 35 40 45
 tgc acc tac ggc gcc ggc gga tcc acc cgc gac cgc acg ttg agt acc 192
 Cys Thr Tyr Gly Ala Gly Gly Ser Thr Arg Asp Arg Thr Leu Ser Thr
 50 55 60
 gtg ctg caa ctg gac ggc gag gtg aag gtg ccg acc gcg ccg cac ctg 240
 Val Leu Gln Leu Asp Gly Glu Val Lys Val Pro Thr Ala Pro His Leu
 65 70 75 80
 tcc tgt gtc ggc gac tcg aaa gcc gag ttg cgc gaa ctg ctc ggc cgc 288
 Ser Cys Val Gly Asp Ser Lys Ala Glu Leu Arg Glu Leu Leu Gly Arg
 85 90 95
 tac cgc gag gcc ggc atc cgc cgc atc gtc gcc ctg cgc ggc gac ctg 336
 Tyr Arg Glu Ala Gly Ile Arg Arg Ile Val Ala Leu Arg Gly Asp Leu
 100 105 110
 ccg tcg ggc atg ggc atg gcc agc ggc gaa ctg cgc tac gcc aac gaa 384
 Pro Ser Gly Met Gly Met Ala Ser Gly Glu Leu Arg Tyr Ala Asn Glu
 115 120 125
 ctg gtg gac ttc atc cgc acc gag acc ggc gac cac ttc cac atc gag 432
 Leu Val Asp Phe Ile Arg Thr Glu Thr Gly Asp His Phe His Ile Glu
 130 135 140
 gtc gcc gcc tat ccg gag gtc cac ccc cag gcg cgc agc ttc gag gat 480
 Val Ala Ala Tyr Pro Glu Val His Pro Gln Ala Arg Ser Phe Glu Asp
 145 150 155 160
 gac ctg gcg aac ttc gtg cgc aag gtg aag gcc ggc gcc agc agc gcc 528
 Asp Leu Ala Asn Phe Val Arg Lys Val Lys Ala Gly Ala Ser Ser Ala
 165 170 175
 atc acc cag tac ttc ttc aac gcc gat gcc tat ttc tac ttc gtc gag 576
 Ile Thr Gln Tyr Phe Phe Asn Ala Asp Ala Tyr Phe Tyr Phe Val Glu
 180 185 190
 cgg gtc gcc aag ctc ggc gtg gac atc ccg gtg gtc ccc ggc atc atg 624
 Arg Val Ala Lys Leu Gly Val Asp Ile Pro Val Val Pro Gly Ile Met
 195 200 205

ccg atc acc aac tac tcc aag ctg gcg cgc ttc tcc gac gcc tgc ggc 672
 Pro Ile Thr Asn Tyr Ser Lys Leu Ala Arg Phe Ser Asp Ala Cys Gly
 210 215 220

gcc gaa ctg ccg cgc tgg atc cgc aag caa ctg gaa gcc tac ggc gac 720
 Ala Glu Leu Pro Arg Trp Ile Arg Lys Gln Leu Glu Ala Tyr Gly Asp
 225 230 235 240

gac agc cgc agc atc cag gcc ttc ggc gag cag gtc atc agc gag atg 768
 Asp Ser Arg Ser Ile Gln Ala Phe Gly Glu Gln Val Ile Ser Glu Met
 245 250 255

tgc gaa cgc ctg ctg gag ggc ggc gca ccg gga ctg cat ttc tat act 816
 Cys Glu Arg Leu Leu Glu Gly Gly Ala Pro Gly Leu His Phe Tyr Thr
 260 265 270

ttg aac cag gcc gat ccg agc ctg gcg atc tgg aag aat ctc cag ctg 864
 Leu Asn Gln Ala Asp Pro Ser Leu Ala Ile Trp Lys Asn Leu Gln Leu
 275 280 285

ca cgc tga 873
 ro Arg
 290

<210> 14
 <211> 290
 <212> PRT
 <213> Pseudomonas aeruginosa

<400> 14
 Val Val Ala Ser Lys Glu Pro Ile Met Ser Gln Ser Glu Arg Arg Phe
 1 5 10 15

Ser Phe Glu Phe Phe Pro Ala Lys Thr Glu Ala Gly His Glu Lys Leu
 20 25 30

Leu Ala Thr Ala Arg Asn Leu Ala Gly Tyr Lys Pro Asp Phe Phe Ser
 35 40 45

ys Thr Tyr Gly Ala Gly Gly Ser Thr Arg Asp Arg Thr Leu Ser Thr
 50 55 60

Val Leu Gln Leu Asp Gly Glu Val Lys Val Pro Thr Ala Pro His Leu
 65 70 75 80

Ser Cys Val Gly Asp Ser Lys Ala Glu Leu Arg Glu Leu Leu Gly Arg
 85 90 95

Tyr Arg Glu Ala Gly Ile Arg Arg Ile Val Ala Leu Arg Gly Asp Leu
 100 105 110

Pro Ser Gly Met Gly Met Ala Ser Gly Glu Leu Arg Tyr Ala Asn Glu
 115 120 125

Leu Val Asp Phe Ile Arg Thr Glu Thr Gly Asp His Phe His Ile Glu
 130 135 140

Val Ala Ala Tyr Pro Glu Val His Pro Gln Ala Arg Ser Phe Glu Asp
 145 150 155 160

Asp Leu Ala Asn Phe Val Arg Lys Val Lys Ala Gly Ala Ser Ser Ala

165										170					175				
Ile	Thr	Gln	Tyr	Phe	Phe	Asn	Ala	Asp	Ala	Tyr	Phe	Tyr	Phe	Val	Glu				
			180					185						190					
Arg	Val	Ala	Lys	Leu	Gly	Val	Asp	Ile	Pro	Val	Val	Pro	Gly	Ile	Met				
		195					200					205							
Pro	Ile	Thr	Asn	Tyr	Ser	Lys	Leu	Ala	Arg	Phe	Ser	Asp	Ala	Cys	Gly				
	210					215					220								
Ala	Glu	Leu	Pro	Arg	Trp	Ile	Arg	Lys	Gln	Leu	Glu	Ala	Tyr	Gly	Asp				
225					230					235					240				
Asp	Ser	Arg	Ser	Ile	Gln	Ala	Phe	Gly	Glu	Gln	Val	Ile	Ser	Glu	Met				
				245					250					255					
Cys	Glu	Arg	Leu	Leu	Glu	Gly	Gly	Ala	Pro	Gly	Leu	His	Phe	Tyr	Thr				
			260					265					270						
Leu	Asn	Gln	Ala	Asp	Pro	Ser	Leu	Ala	Ile	Trp	Lys	Asn	Leu	Gln	Leu				
		275					280					285							
Pro	Arg																		
	290																		

<210> 15
 <211> 828
 <212> DNA
 <213> Xylella almond

<220>
 <221> CDS
 <222> (1) .. (825)
 <223> RXFX00359

<400> 15																
atg	att	cca	atc	agc	ttc	gag	ttt	tat	cca	ccc	aaa	aac	gac	gat	caa	48
Met	Ile	Pro	Ile	Ser	Phe	Glu	Phe	Tyr	Pro	Pro	Lys	Asn	Asp	Asp	Gln	
1				5					10					15		
cgc	gca	cag	ttg	gac	agg	aca	gca	aac	cgg	cta	cgc	gca	ttc	gca	cca	96
Arg	Ala	Gln	Leu	Asp	Arg	Thr	Ala	Asn	Arg	Leu	Arg	Ala	Phe	Ala	Pro	
			20					25					30			
gaa	tac	gtc	tcc	tgc	acc	ttc	ggc	gcc	ggt	ggc	tcc	aca	ctc	agt	tac	144
Glu	Tyr	Val	Ser	Cys	Thr	Phe	Gly	Ala	Gly	Gly	Ser	Thr	Leu	Ser	Tyr	
		35					40					45				
acc	tca	gaa	aca	gtg	cgc	cat	ctc	agc	caa	cac	cac	ggc	ttt	gac	gcc	192
Thr	Ser	Glu	Thr	Val	Arg	His	Leu	Ser	Gln	His	His	Gly	Phe	Asp	Ala	
		50				55					60					
gca	ccg	cat	ctg	tcc	tgt	gtg	ggc	ggc	agt	cgc	caa	gaa	atc	cgc	gaa	240
Ala	Pro	His	Leu	Ser	Cys	Val	Gly	Gly	Ser	Arg	Gln	Glu	Ile	Arg	Glu	
65					70				75					80		
ctt	ctc	aaa	ctg	tac	cgc	gcg	att	ggc	tgc	caa	cgc	atc	gtg	gcg	cta	288
Leu	Leu	Lys	Leu	Tyr	Arg	Ala	Ile	Gly	Cys	Gln	Arg	Ile	Val	Ala	Leu	
			85					90						95		

cgc ggc gat ctc ccc tcg ggc atg ggc cac ccc ggc gac ctc cgc tac 336
 Arg Gly Asp Leu Pro Ser Gly Met Gly His Pro Gly Asp Leu Arg Tyr
 100 105 110

gca gct gac ctg att acc ttc atc cgt acc gag cat ggc gat cac ttc 384
 Ala Ala Asp Leu Ile Thr Phe Ile Arg Thr Glu His Gly Asp His Phe
 115 120 125

cac cta gag atc ggc gca tac ccg gaa acc cac cca caa gcc agc aac 432
 His Leu Glu Ile Gly Ala Tyr Pro Glu Thr His Pro Gln Ala Ser Asn
 130 135 140

aca ctg aac gac ctt cac tat ttc aaa gcc aaa gcc gat gca ggc gcc 480
 Thr Leu Asn Asp Leu His Tyr Phe Lys Ala Lys Ala Asp Ala Gly Ala
 145 150 155 160

gat gcg gca atc act caa tac ttt tat aac cca gac gcc tat ttc cac 528
 Asp Ala Ala Ile Thr Gln Tyr Phe Tyr Asn Pro Asp Ala Tyr Phe His
 165 170 175

ttc gtc gac gca gtg cag cgc ctg ggc gtc acc atc ccc att gtt gcc 576
 Phe Val Asp Ala Val Gln Arg Leu Gly Val Thr Ile Pro Ile Val Ala
 180 185 190

gga gtc atg ccc atc tcc aac ttt gac cag ttg cgc cat ttc tcc gaa 624
 Gly Val Met Pro Ile Ser Asn Phe Asp Gln Leu Arg His Phe Ser Glu
 195 200 205

caa tgc ggc gcc gaa ata ccc cgc tgg att aca aaa aaa atg cag gct 672
 Gln Cys Gly Ala Glu Ile Pro Arg Trp Ile Thr Lys Lys Met Gln Ala
 210 215 220

tac ggc gac gac acc aaa tcg ata cgc gcg ttc ggt gcc gac gtc gtg 720
 Tyr Gly Asp Asp Thr Lys Ser Ile Arg Ala Phe Gly Ala Asp Val Val
 225 230 235 240

acc gca tta tgt gag cgg cta atc gct ggc ggc gca ccg ggg ctg cac 768
 Thr Ala Leu Cys Glu Arg Leu Ile Ala Gly Gly Ala Pro Gly Leu His
 245 250 255

ttc tac acg ctc aac cta gcc aaa cca agc acc caa gtg ctg caa cgc 816
 Phe Tyr Thr Leu Asn Leu Ala Lys Pro Ser Thr Gln Val Leu Gln Arg
 260 265 270

tta ggc tat tga 828
 Leu Gly Tyr
 275

<210> 16
 <211> 275
 <212> PRT
 <213> Xylella almond

<400> 16
 Met Ile Pro Ile Ser Phe Glu Phe Tyr Pro Pro Lys Asn Asp Asp Gln
 1 5 10 15
 Arg Ala Gln Leu Asp Arg Thr Ala Asn Arg Leu Arg Ala Phe Ala Pro
 20 25 30
 Glu Tyr Val Ser Cys Thr Phe Gly Ala Gly Gly Ser Thr Leu Ser Tyr
 35 40 45

Thr Ser Glu Thr Val Arg His Leu Ser Gln His His Gly Phe Asp Ala

50

55

60

Ala Pro His Leu Ser Cys Val Gly Gly Ser Arg Gln Glu Ile Arg Glu
65 70 75 80

Leu Leu Lys Leu Tyr Arg Ala Ile Gly Cys Gln Arg Ile Val Ala Leu
85 90 95

Arg Gly Asp Leu Pro Ser Gly Met Gly His Pro Gly Asp Leu Arg Tyr
100 105 110

Ala Ala Asp Leu Ile Thr Phe Ile Arg Thr Glu His Gly Asp His Phe
115 120 125

His Leu Glu Ile Gly Ala Tyr Pro Glu Thr His Pro Gln Ala Ser Asn
130 135 140

Thr Leu Asn Asp Leu His Tyr Phe Lys Ala Lys Ala Asp Ala Gly Ala
145 150 155 160

Asp Ala Ala Ile Thr Gln Tyr Phe Tyr Asn Pro Asp Ala Tyr Phe His
165 170 175

Phe Val Asp Ala Val Gln Arg Leu Gly Val Thr Ile Pro Ile Val Ala
180 185 190

Gly Val Met Pro Ile Ser Asn Phe Asp Gln Leu Arg His Phe Ser Glu
195 200 205

Gln Cys Gly Ala Glu Ile Pro Arg Trp Ile Thr Lys Lys Met Gln Ala
210 215 220

Tyr Gly Asp Asp Thr Lys Ser Ile Arg Ala Phe Gly Ala Asp Val Val
225 230 235 240

Thr Ala Leu Cys Glu Arg Leu Ile Ala Gly Gly Ala Pro Gly Leu His
245 250 255

Phe Tyr Thr Leu Asn Leu Ala Lys Pro Ser Thr Gln Val Leu Gln Arg
260 265 270

Leu Gly Tyr
275

<210> 17

<211> 828

<212> DNA

<213> Xylella oleander

<220>

<221> CDS

<222> (1) .. (825)

<223> RXFY01676

<400> 17

atg att cca atc agc ttc gag ttt tat cca ccc aaa aac gac gat caa 48
Met Ile Pro Ile Ser Phe Glu Phe Tyr Pro Pro Lys Asn Asp Asp Gln
1 5 10 15

cgc gca cag ttg gac agg aca gca aac cgg cta cgc gca ttc gca cca	96
Arg Ala Gln Leu Asp Arg Thr Ala Asn Arg Leu Arg Ala Phe Ala Pro	
20 25 30	
gaa tac gtc tcc tgc acc ttc ggc gcc ggc ggc tcc aca ctc agt tac	144
Glu Tyr Val Ser Cys Thr Phe Gly Ala Gly Gly Ser Thr Leu Ser Tyr	
35 40 45	
acc tca gaa aca gtg cgc cat ctc agt caa cac cac ggc ttt gac acc	192
Thr Ser Glu Thr Val Arg His Leu Ser Gln His His Gly Phe Asp Thr	
50 55 60	
gca ccg cat ctg tcc tgt gtg ggc gcc agt cgc caa gaa atc cgc gaa	240
Ala Pro His Leu Ser Cys Val Gly Gly Ser Arg Gln Glu Ile Arg Glu	
65 70 75 80	
ctt ctc aaa ctg tac cgc gcg att ggc tgc caa cgc atc gtg gcg cta	288
Leu Leu Lys Leu Tyr Arg Ala Ile Gly Cys Gln Arg Ile Val Ala Leu	
85 90 95	
cgc ggc gat ctc ccc tcg ggc atg ggc cac ccc ggc gac ctc cgc tac	336
Arg Gly Asp Leu Pro Ser Gly Met Gly His Pro Gly Asp Leu Arg Tyr	
100 105 110	
gca gct gac ctg att acc ttc atc cgt gcc gag cat ggc gat cac ttc	384
Ala Ala Asp Leu Ile Thr Phe Ile Arg Ala Glu His Gly Asp His Phe	
115 120 125	
cac cta gag atc ggc gca tac ccg gaa acc cac cca caa gcc agc aac	432
His Leu Glu Ile Gly Ala Tyr Pro Glu Thr His Pro Gln Ala Ser Asn	
130 135 140	
aca ctg aac gac ctt cac tat ttc aaa gcc aaa gcc gat gca ggc gcc	480
Thr Leu Asn Asp Leu His Tyr Phe Lys Ala Lys Ala Asp Ala Gly Ala	
145 150 155 160	
gat gcg gca atc act caa tac ttt tac aac cca gac gcc tat ttc cac	528
Asp Ala Ala Ile Thr Gln Tyr Phe Tyr Asn Pro Asp Ala Tyr Phe His	
165 170 175	
ttc gtc gac gca gtg cag cgc ctg ggc gtc acc atc ccc att gtt gcc	576
Phe Val Asp Ala Val Gln Arg Leu Gly Val Thr Ile Pro Ile Val Ala	
180 185 190	
gga gtc atg ccc atc tcc aac ttt gac cag ttg cgc cat ttc tcc gaa	624
Gly Val Met Pro Ile Ser Asn Phe Asp Gln Leu Arg His Phe Ser Glu	
195 200 205	
caa tgc ggc gcc gaa ata ccc cgc tgg att aca aaa aaa atg cag gct	672
Gln Cys Gly Ala Glu Ile Pro Arg Trp Ile Thr Lys Lys Met Gln Ala	
210 215 220	
tac ggc gat gac acc aaa tcg ata cgc gcg ttc ggt gcc gac gtc gtg	720
Tyr Gly Asp Asp Thr Lys Ser Ile Arg Ala Phe Gly Ala Asp Val Val	
225 230 235 240	
acc gca cta tgt gag cgg cta atc gct ggc gcc gca ccg ggg ctg cac	768
Thr Ala Leu Cys Glu Arg Leu Ile Ala Gly Gly Ala Pro Gly Leu His	
245 250 255	

ttc tac acg ctc aac cta gcc aaa cca agc acc caa gtg ctg caa cgc 816
 Phe Tyr Thr Leu Asn Leu Ala Lys Pro Ser Thr Gln Val Leu Gln Arg
 260 265 270

tta ggc tat tga 828
 Leu Gly Tyr
 275

<210> 18

<211> 275

<212> PRT

<213> Xylella oleander

<400> 18

Met Ile Pro Ile Ser Phe Glu Phe Tyr Pro Pro Lys Asn Asp Asp Gln
 1 5 10 15

Arg Ala Gln Leu Asp Arg Thr Ala Asn Arg Leu Arg Ala Phe Ala Pro
 20 25 30

Glu Tyr Val Ser Cys Thr Phe Gly Ala Gly Gly Ser Thr Leu Ser Tyr
 35 40 45

Thr Ser Glu Thr Val Arg His Leu Ser Gln His His Gly Phe Asp Thr
 50 55 60

Ala Pro His Leu Ser Cys Val Gly Gly Ser Arg Gln Glu Ile Arg Glu
 65 70 75 80

Leu Leu Lys Leu Tyr Arg Ala Ile Gly Cys Gln Arg Ile Val Ala Leu
 85 90 95

Arg Gly Asp Leu Pro Ser Gly Met Gly His Pro Gly Asp Leu Arg Tyr
 100 105 110

Ala Ala Asp Leu Ile Thr Phe Ile Arg Ala Glu His Gly Asp His Phe
 115 120 125

His Leu Glu Ile Gly Ala Tyr Pro Glu Thr His Pro Gln Ala Ser Asn
 130 135 140

Thr Leu Asn Asp Leu His Tyr Phe Lys Ala Lys Ala Asp Ala Gly Ala
 145 150 155 160

Asp Ala Ala Ile Thr Gln Tyr Phe Tyr Asn Pro Asp Ala Tyr Phe His
 165 170 175

Phe Val Asp Ala Val Gln Arg Leu Gly Val Thr Ile Pro Ile Val Ala
 180 185 190

Gly Val Met Pro Ile Ser Asn Phe Asp Gln Leu Arg His Phe Ser Glu
 195 200 205

Gln Cys Gly Ala Glu Ile Pro Arg Trp Ile Thr Lys Lys Met Gln Ala
 210 215 220

Tyr Gly Asp Asp Thr Lys Ser Ile Arg Ala Phe Gly Ala Asp Val Val
 225 230 235 240

Thr Ala Leu Cys Glu Arg Leu Ile Ala Gly Gly Ala Pro Gly Leu His
 245 250 255

Phe Tyr Thr Leu Asn Leu Ala Lys Pro Ser Thr Gln Val Leu Gln Arg
 260 265 270

Leu Gly Tyr
 275

<210> 19

<211> 846

<212> DNA

<213> *Pseudomonas fluorescens*

<220>

<221> CDS

<222> (1) .. (843)

<223> RPU04845

<400> 19

ctg tcc caa gac cgt cgc tac agc ttc gag ttc ttc ccg acc aag acc	48
t Ser Gln Asp Arg Arg Tyr Ser Phe Glu Phe Phe Pro Thr Lys Thr	
1 5 10 15	
gat gct ggg cat gaa aaa ctg ctc gcc act gcc cgt cag ctg gcc acc	96
Asp Ala Gly His Glu Lys Leu Leu Ala Thr Ala Arg Gln Leu Ala Thr	
20 25 30	
tat aag cct gac ttc ttt tcc tgc acc tac ggc gct ggc ggt tcg acc	144
Tyr Lys Pro Asp Phe Phe Ser Cys Thr Tyr Gly Ala Gly Gly Ser Thr	
35 40 45	
cgt gac cgc acg ctg aac acc gtt ctg cag ctg gaa agc gaa gtc aaa	192
Arg Asp Arg Thr Leu Asn Thr Val Leu Gln Leu Glu Ser Glu Val Lys	
50 55 60	
atc ccc gcc gca ccg cac ctg tcg tgc gtc ggc gac agc aag gac gac	240
Ile Pro Ala Ala Pro His Leu Ser Cys Val Gly Asp Ser Lys Asp Asp	
65 70 75 80	
ctg cgc ggc ctg ctg aac gag tac aag gcc gcc ggc atc aag cgc atc	288
Leu Arg Gly Leu Leu Asn Glu Tyr Lys Ala Ala Gly Ile Lys Arg Ile	
85 90 95	
gtc gcc ctg cgc ggt gac ctg ccg tcc ggc atg ggc atg acc agc ggc	336
Val Ala Leu Arg Gly Asp Leu Pro Ser Gly Met Gly Met Thr Ser Gly	
100 105 110	
gag ctg cgt cac gcc aat gaa ctg gtt gaa ttc att cgt gaa gaa acc	384
Glu Leu Arg His Ala Asn Glu Leu Val Glu Phe Ile Arg Glu Glu Thr	
115 120 125	
ggc aat cat ttc cac atc gaa gtc gcc gcc tac ccg gag atg cat ccg	432
Gly Asn His Phe His Ile Glu Val Ala Ala Tyr Pro Glu Met His Pro	
130 135 140	
caa gcg cgc aac tac gaa gac gat ctc gcc aac ttc gtg cgc aag gcc	480
Gln Ala Arg Asn Tyr Glu Asp Asp Leu Ala Asn Phe Val Arg Lys Ala	
145 150 155 160	
cgt gcc ggc gcc gac agc gcg atc acc cag tac ttc ttc aac gcc gac	528
Arg Ala Gly Ala Asp Ser Ala Ile Thr Gln Tyr Phe Phe Asn Ala Asp	
165 170 175	

agc tac ttc tac ttc gtc gac cgt ttg cag gcg ctg ggc gtg gac att 576
 Ser Tyr Phe Tyr Phe Val Asp Arg Leu Gln Ala Leu Gly Val Asp Ile
 180 185 190

ccg gtg gta ccg ggg atc atg ccg atc acc aac tac agc aaa ctc gcg 624
 Pro Val Val Pro Gly Ile Met Pro Ile Thr Asn Tyr Ser Lys Leu Ala
 195 200 205

cgc ttc tcc gat gcc tgc ggt gcg gaa atc ccg cgc tgg atc cgc aag 672
 Arg Phe Ser Asp Ala Cys Gly Ala Glu Ile Pro Arg Trp Ile Arg Lys
 210 215 220

cag ctg gaa gcc tac ggc gat gac agc caa agc att cag cgc ttt ggc 720
 Gln Leu Glu Ala Tyr Gly Asp Asp Ser Gln Ser Ile Gln Arg Phe Gly
 225 230 235 240

gaa caa gtc gtc acg gaa atg tgc gaa cgc ctg ctg caa ggc ggc gcg 768
 Glu Gln Val Val Thr Glu Met Cys Glu Arg Leu Leu Gln Gly Gly Ala
 245 250 255

ccg ggc ctg cac ttc tat tcc atg aac cag gcc gaa cca agc ctg gcg 816
 Pro Gly Leu His Phe Tyr Ser Met Asn Gln Ala Glu Pro Ser Leu Ala
 260 265 270

atc tgg aac aac ctg aag ctg ccg cgc taa 846
 Ile Trp Asn Asn Leu Lys Leu Pro Arg
 275 280

<210> 20

<211> 281

<212> PRT

<213> *Pseudomonas fluorescens*

<400> 20

Met Ser Gln Asp Arg Arg Tyr Ser Phe Glu Phe Phe Pro Thr Lys Thr
 1 5 10 15

Asp Ala Gly His Glu Lys Leu Leu Ala Thr Ala Arg Gln Leu Ala Thr
 20 25 30

Tyr Lys Pro Asp Phe Phe Ser Cys Thr Tyr Gly Ala Gly Gly Ser Thr
 35 40 45

Arg Asp Arg Thr Leu Asn Thr Val Leu Gln Leu Glu Ser Glu Val Lys
 50 55 60

Ile Pro Ala Ala Pro His Leu Ser Cys Val Gly Asp Ser Lys Asp Asp
 65 70 75 80

Leu Arg Gly Leu Leu Asn Glu Tyr Lys Ala Ala Gly Ile Lys Arg Ile
 85 90 95

Val Ala Leu Arg Gly Asp Leu Pro Ser Gly Met Gly Met Thr Ser Gly
 100 105 110

Glu Leu Arg His Ala Asn Glu Leu Val Glu Phe Ile Arg Glu Glu Thr
 115 120 125

Gly Asn His Phe His Ile Glu Val Ala Ala Tyr Pro Glu Met His Pro
 130 135 140

Gln Ala Arg Asn Tyr Glu Asp Asp Leu Ala Asn Phe Val Arg Lys Ala
145 150 155 160

Arg Ala Gly Ala Asp Ser Ala Ile Thr Gln Tyr Phe Phe Asn Ala Asp
165 170 175

Ser Tyr Phe Tyr Phe Val Asp Arg Leu Gln Ala Leu Gly Val Asp Ile
180 185 190

Pro Val Val Pro Gly Ile Met Pro Ile Thr Asn Tyr Ser Lys Leu Ala
195 200 205

Arg Phe Ser Asp Ala Cys Gly Ala Glu Ile Pro Arg Trp Ile Arg Lys
210 215 220

Gln Leu Glu Ala Tyr Gly Asp Asp Ser Gln Ser Ile Gln Arg Phe Gly
225 230 235 240

Glu Gln Val Val Thr Glu Met Cys Glu Arg Leu Leu Gln Gly Gly Ala
245 250 255

Pro Gly Leu His Phe Tyr Ser Met Asn Gln Ala Glu Pro Ser Leu Ala
260 265 270

Ile Trp Asn Asn Leu Lys Leu Pro Arg
275 280

<210> 21

<211> 1812

<212> DNA

<213> Schizosaccharomyces pombe

<220>

<221> CDS

<222> (1) .. (1809)

<223> RSO01645

<400> 21

atg aaa ata agt gac aaa tta ctt cac ccg gat tgg aag gaa aaa gtt 48
Met Lys Ile Ser Asp Lys Leu Leu His Pro Asp Trp Lys Glu Lys Val
1 5 10 15

act tac agt tat gaa ttt ttt cct cca aaa acg agc act ggt gtc caa 96
Thr Tyr Ser Tyr Glu Phe Phe Pro Pro Lys Thr Ser Thr Gly Val Gln
20 25 30

aat ctt tac aat cgt ata gat cgc atg aag act tgg ggt cgt ccc atg 144
Asn Leu Tyr Asn Arg Ile Asp Arg Met Lys Thr Trp Gly Arg Pro Met
35 40 45

ttt gtc gat gtg act tgg ggt gct ggt ggt act tct tca gaa ctg act 192
Phe Val Asp Val Thr Trp Gly Ala Gly Gly Thr Ser Ser Glu Leu Thr
50 55 60

cct gga atc gtt aat gta att caa aca gat ttt gaa gtg gat act tgc 240
Pro Gly Ile Val Asn Val Ile Gln Thr Asp Phe Glu Val Asp Thr Cys
65 70 75 80

atg cat ttg act tgt acg aac atg tcc aca gaa atg att gac gca gct 288
Met His Leu Thr Cys Thr Asn Met Ser Thr Glu Met Ile Asp Ala Ala
85 90 95

ttg aaa cgg gct cat gaa aca ggg tgt cgt aac ata ttg gcc ctt aga	336
Leu Lys Arg Ala His Glu Thr Gly Cys Arg Asn Ile Leu Ala Leu Arg	
100 105 110	
ggt gat cct gtt aaa gat aca gac tgg act gaa ggc gaa agt gga ttc	384
Gly Asp Pro Val Lys Asp Thr Asp Trp Thr Glu Gly Glu Ser Gly Phe	
115 120 125	
cgg tat gct tca gac tta gtt aga tat att cgc aca cat tat aat gat	432
Arg Tyr Ala Ser Asp Leu Val Arg Tyr Ile Arg Thr His Tyr Asn Asp	
130 135 140	
gaa ttc tgt att ggt gta gct ggc tat cca gaa gga tat tca cca gat	480
Glu Phe Cys Ile Gly Val Ala Gly Tyr Pro Glu Gly Tyr Ser Pro Asp	
145 150 155 160	
gat gac att gat gaa agc ata aag cat ctg aaa tta aaa gtc gat gaa	528
Asp Asp Ile Asp Glu Ser Ile Lys His Leu Lys Leu Lys Val Asp Glu	
165 170 175	
ggt gct gat ttt atc gtt act caa atg ttt tat gat gta gac aat ttt	576
Gly Ala Asp Phe Ile Val Thr Gln Met Phe Tyr Asp Val Asp Asn Phe	
180 185 190	
atc gca tgg gtc gat aaa gtg cgt gca gca gga ata aat atc cct ata	624
Ile Ala Trp Val Asp Lys Val Arg Ala Ala Gly Ile Asn Ile Pro Ile	
195 200 205	
ttt ccg ggc att atg cct att cag gca tgg gat tcc ttt att cgg aga	672
Phe Pro Gly Ile Met Pro Ile Gln Ala Trp Asp Ser Phe Ile Arg Arg	
210 215 220	
gcg aaa tgg agc ggt gtt aaa att ccc cag cat ttt atg gat act cta	720
Ala Lys Trp Ser Gly Val Lys Ile Pro Gln His Phe Met Asp Thr Leu	
225 230 235 240	
gtc cca gtt aaa gac gat gat gaa gga gtc cgt gag cgt ggt gtt gag	768
Val Pro Val Lys Asp Asp Asp Glu Gly Val Arg Glu Arg Gly Val Glu	
245 250 255	
ctc ata gtc gaa atg tgc cgt aag ctt ata gct agt ggc att acg aga	816
Leu Ile Val Glu Met Cys Arg Lys Leu Ile Ala Ser Gly Ile Thr Arg	
260 265 270	
ctt cat ttt tac act atg aat tta gaa aag gcc gtt aaa atg att att	864
Leu His Phe Tyr Thr Met Asn Leu Glu Lys Ala Val Lys Met Ile Ile	
275 280 285	
gaa cga tta ggt tta tta gat gaa aac ttg gct cct ata gtg gat act	912
Glu Arg Leu Gly Leu Leu Asp Glu Asn Leu Ala Pro Ile Val Asp Thr	
290 295 300	
aat aac gtc gag tta acc aat gct tcc agt caa gat cgt cgg ata aat	960
Asn Asn Val Glu Leu Thr Asn Ala Ser Ser Gln Asp Arg Arg Ile Asn	
305 310 315 320	
gaa ggt gta cgg ccc att ttc tgg cgc act cgt aat gaa agt tat gtc	1008
Glu Gly Val Arg Pro Ile Phe Trp Arg Thr Arg Asn Glu Ser Tyr Val	
325 330 335	
tcg cgt act gat cag tgg gac gaa tta ccg cat ggt cgt tgg ggt gac	1056
Ser Arg Thr Asp Gln Trp Asp Glu Leu Pro His Gly Arg Trp Gly Asp	
340 345 350	

tct cgt agc cct gct ttt ggc gaa ttt gat gct att aga tat ggt ctt	1104
Ser Arg Ser Pro Ala Phe Gly Glu Phe Asp Ala Ile Arg Tyr Gly Leu	
355 360 365	
cgt atg tct ccc aag gag atc aca aca tcg tgg ggg tct cct aaa tct	1152
Arg Met Ser Pro Lys Glu Ile Thr Thr Ser Trp Gly Ser Pro Lys Ser	
370 375 380	
tac tcg gaa atc ggc gat ttg ttt gcc agg tac tgt gaa aaa aag att	1200
Tyr Ser Glu Ile Gly Asp Leu Phe Ala Arg Tyr Cys Glu Lys Lys Ile	
385 390 395 400	
agc tcc ctc cct tgg agt gat ctt ccc ata tcc gat gaa gcc gac ttg	1248
Ser Ser Leu Pro Trp Ser Asp Leu Pro Ile Ser Asp Glu Ala Asp Leu	
405 410 415	
att cgg gat caa ctt cta agt atg aat aga aac gct ttc ctt act ata	1296
Ile Arg Asp Gln Leu Leu Ser Met Asn Arg Asn Ala Phe Leu Thr Ile	
420 425 430	
t tct caa cct gct ctt aac ggc gaa aag agt tca cat cct gtt ttt	1344
sn Ser Gln Pro Ala Leu Asn Gly Glu Lys Ser Ser His Pro Val Phe	
435 440 445	
gga tgg gga cca cct aat ggt tat gtt ttc caa aaa cca tac gtt gag	1392
Gly Trp Gly Pro Pro Asn Gly Tyr Val Phe Gln Lys Pro Tyr Val Glu	
450 455 460	
ttt ttc gtt cac ccc tca ctc ttg aat gaa ctc aaa gaa acc gtt aaa	1440
Phe Phe Val His Pro Ser Leu Leu Asn Glu Leu Lys Glu Thr Val Lys	
465 470 475 480	
aag ctt aat tca gtt tcc tac ttt gtt aca aac aag aat gga gac ttg	1488
Lys Leu Asn Ser Val Ser Tyr Phe Val Thr Asn Lys Asn Gly Asp Leu	
485 490 495	
gat acc aac tca caa tat gag att cca aat gcg gtt aca tgg ggt gtt	1536
Asp Thr Asn Ser Gln Tyr Glu Ile Pro Asn Ala Val Thr Trp Gly Val	
500 505 510	
tc cct aat cgt gag att atc caa cct act att gtc gag tca acc tct	1584
ne Pro Asn Arg Glu Ile Ile Gln Pro Thr Ile Val Glu Ser Thr Ser	
515 520 525	
ttt ctt gct tgg aaa gat gaa gcc tat tca ttg ggc atg gaa tgg gct	1632
Phe Leu Ala Trp Lys Asp Glu Ala Tyr Ser Leu Gly Met Glu Trp Ala	
530 535 540	
aat gca tat agc cct gat tca att tct cgt aaa ctt ttg gtt tct atg	1680
Asn Ala Tyr Ser Pro Asp Ser Ile Ser Arg Lys Leu Leu Val Ser Met	
545 550 555 560	
atg aag gaa tgg ttc ctt tgt gtc ata gtt gat aac gat ttt caa aat	1728
Met Lys Glu Trp Phe Leu Cys Val Ile Val Asp Asn Asp Phe Gln Asn	
565 570 575	
ggg caa tct ttg ttt gat gtt ttt aac aaa atg aga tct tta aaa gac	1776
Gly Gln Ser Leu Phe Asp Val Phe Asn Lys Met Arg Ser Leu Lys Asp	
580 585 590	
atc cat cct gag cta tat tat gca aat gca tca taa	1812
Ile His Pro Glu Leu Tyr Tyr Ala Asn Ala Ser	

595

600

<210> 22
 <211> 603
 <212> PRT
 <213> Schizosaccharomyces pombe

<400> 22

Met Lys Ile Ser Asp Lys Leu Leu His Pro Asp Trp Lys Glu Lys Val
 1 5 10 15

Thr Tyr Ser Tyr Glu Phe Phe Pro Pro Lys Thr Ser Thr Gly Val Gln
 20 25 30

Asn Leu Tyr Asn Arg Ile Asp Arg Met Lys Thr Trp Gly Arg Pro Met
 35 40 45

Phe Val Asp Val Thr Trp Gly Ala Gly Gly Thr Ser Ser Glu Leu Thr
 50 55 60

Pro Gly Ile Val Asn Val Ile Gln Thr Asp Phe Glu Val Asp Thr Cys
 65 70 75 80

Met His Leu Thr Cys Thr Asn Met Ser Thr Glu Met Ile Asp Ala Ala
 85 90 95

Leu Lys Arg Ala His Glu Thr Gly Cys Arg Asn Ile Leu Ala Leu Arg
 100 105 110

Gly Asp Pro Val Lys Asp Thr Asp Trp Thr Glu Gly Glu Ser Gly Phe
 115 120 125

Arg Tyr Ala Ser Asp Leu Val Arg Tyr Ile Arg Thr His Tyr Asn Asp
 130 135 140

Glu Phe Cys Ile Gly Val Ala Gly Tyr Pro Glu Gly Tyr Ser Pro Asp
 145 150 155 160

Asp Asp Ile Asp Glu Ser Ile Lys His Leu Lys Leu Lys Val Asp Glu
 165 170 175

Gly Ala Asp Phe Ile Val Thr Gln Met Phe Tyr Asp Val Asp Asn Phe
 180 185 190

Ile Ala Trp Val Asp Lys Val Arg Ala Ala Gly Ile Asn Ile Pro Ile
 195 200 205

Phe Pro Gly Ile Met Pro Ile Gln Ala Trp Asp Ser Phe Ile Arg Arg
 210 215 220

Ala Lys Trp Ser Gly Val Lys Ile Pro Gln His Phe Met Asp Thr Leu
 225 230 235 240

Val Pro Val Lys Asp Asp Asp Glu Gly Val Arg Glu Arg Gly Val Glu
 245 250 255

Leu Ile Val Glu Met Cys Arg Lys Leu Ile Ala Ser Gly Ile Thr Arg
 260 265 270

Leu His Phe Tyr Thr Met Asn Leu Glu Lys Ala Val Lys Met Ile Ile
 275 280 285

Glu Arg Leu Gly Leu Leu Asp Glu Asn Leu Ala Pro Ile Val Asp Thr
 290 295 300
 Asn Asn Val Glu Leu Thr Asn Ala Ser Ser Gln Asp Arg Arg Ile Asn
 305 310 315 320
 Glu Gly Val Arg Pro Ile Phe Trp Arg Thr Arg Asn Glu Ser Tyr Val
 325 330 335
 Ser Arg Thr Asp Gln Trp Asp Glu Leu Pro His Gly Arg Trp Gly Asp
 340 345 350
 Ser Arg Ser Pro Ala Phe Gly Glu Phe Asp Ala Ile Arg Tyr Gly Leu
 355 360 365
 Arg Met Ser Pro Lys Glu Ile Thr Thr Ser Trp Gly Ser Pro Lys Ser
 370 375 380
 Tyr Ser Glu Ile Gly Asp Leu Phe Ala Arg Tyr Cys Glu Lys Lys Ile
 385 390 395 400
 Ser Ser Leu Pro Trp Ser Asp Leu Pro Ile Ser Asp Glu Ala Asp Leu
 405 410 415
 Ile Arg Asp Gln Leu Leu Ser Met Asn Arg Asn Ala Phe Leu Thr Ile
 420 425 430
 Asn Ser Gln Pro Ala Leu Asn Gly Glu Lys Ser Ser His Pro Val Phe
 435 440 445
 Gly Trp Gly Pro Pro Asn Gly Tyr Val Phe Gln Lys Pro Tyr Val Glu
 450 455 460
 Phe Phe Val His Pro Ser Leu Leu Asn Glu Leu Lys Glu Thr Val Lys
 465 470 475 480
 Lys Leu Asn Ser Val Ser Tyr Phe Val Thr Asn Lys Asn Gly Asp Leu
 485 490 495
 Asp Thr Asn Ser Gln Tyr Glu Ile Pro Asn Ala Val Thr Trp Gly Val
 500 505 510
 Phe Pro Asn Arg Glu Ile Ile Gln Pro Thr Ile Val Glu Ser Thr Ser
 515 520 525
 Phe Leu Ala Trp Lys Asp Glu Ala Tyr Ser Leu Gly Met Glu Trp Ala
 530 535 540
 Asn Ala Tyr Ser Pro Asp Ser Ile Ser Arg Lys Leu Leu Val Ser Met
 545 550 555 560
 Met Lys Glu Trp Phe Leu Cys Val Ile Val Asp Asn Asp Phe Gln Asn
 565 570 575
 Gly Gln Ser Leu Phe Asp Val Phe Asn Lys Met Arg Ser Leu Lys Asp
 580 585 590
 Ile His Pro Glu Leu Tyr Tyr Ala Asn Ala Ser
 595 600

<210> 23

<211> 1800

M/43126

MetF

<212> DNA

<213> *Saccharomyces cerevisiae*

<220>

<221> CDS

<222> (1)..(1797)

<223> RSC08323

<400> 23

atg aag atc aca gaa aaa tta gag caa cat aga cag acc tct ggc aag	48
Met Lys Ile Thr Glu Lys Leu Glu Gln His Arg Gln Thr Ser Gly Lys	
1 5 10 15	
ccc act tac tca ttc gag tac ttc gtc ccg aag act aca caa ggt gta	96
Pro Thr Tyr Ser Phe Glu Tyr Phe Val Pro Lys Thr Thr Gln Gly Val	
20 25 30	
cag aac ctg tat gac cgg atg gac cgg atg tac gag gct tct ttg ccc	144
Gln Asn Leu Tyr Asp Arg Met Asp Arg Met Tyr Glu Ala Ser Leu Pro	
35 40 45	
aaa ttt att gac atc acc tgg aat gca ggc ggt gga cgg ttg tca cat	192
Gln Phe Ile Asp Ile Thr Trp Asn Ala Gly Gly Gly Arg Leu Ser His	
50 55 60	
ctg tcc acg gac ttg gtt gcg aca gcg cag tct gtg ctt ggt ttg gaa	240
Leu Ser Thr Asp Leu Val Ala Thr Ala Gln Ser Val Leu Gly Leu Glu	
65 70 75 80	
acg tgc atg cac ctt acc tgc acc aat atg ccc att tcg atg att gac	288
Thr Cys Met His Leu Thr Cys Thr Asn Met Pro Ile Ser Met Ile Asp	
85 90 95	
gac gct tta gaa aac gct tat cac tcc ggt tgc cag aac atc cta gcg	336
Asp Ala Leu Glu Asn Ala Tyr His Ser Gly Cys Gln Asn Ile Leu Ala	
100 105 110	
ctg aga gga gat cct cct agg gac gca gaa aac tgg act ccc gtt gaa	384
Leu Arg Gly Asp Pro Pro Arg Asp Ala Glu Asn Trp Thr Pro Val Glu	
115 120 125	
gtg ggc ttc cag tat gcc aag gac ttg att aag tat atc aag tcc aag	432
Gly Gly Phe Gln Tyr Ala Lys Asp Leu Ile Lys Tyr Ile Lys Ser Lys	
130 135 140	
tac ggt gac cat ttc gct atc ggc gtt gcc ggc tac ccg gag tgc cat	480
Tyr Gly Asp His Phe Ala Ile Gly Val Ala Gly Tyr Pro Glu Cys His	
145 150 155 160	
ccg gag ttg cct aac aaa gac gtg aag ctt gat ctc gag tat ttg agc	528
Pro Glu Leu Pro Asn Lys Asp Val Lys Leu Asp Leu Glu Tyr Leu Ser	
165 170 175	
aga aga tcg acc ggc ggc gac ttc atc atc act cag atg ttt tac gat	576
Arg Arg Ser Thr Gly Gly Asp Phe Ile Ile Thr Gln Met Phe Tyr Asp	
180 185 190	
gtt gat aat tta ctc aac tgg tgt tcc caa gtt aga gct gcg ggc atg	624
Val Asp Asn Leu Leu Asn Trp Cys Ser Gln Val Arg Ala Ala Gly Met	
195 200 205	
gac gtg ccc att att ccc ggg atc atg ccg atc act acc tac gcg gcc	672

Asp Val Pro Ile Ile Pro Gly Ile Met Pro Ile Thr Thr Tyr Ala Ala	
210 215 220	
ttc ttg aga agg atc caa tgg ggc caa atc tcc atc cct caa cat ttc	720
Phe Leu Arg Arg Ile Gln Trp Gly Gln Ile Ser Ile Pro Gln His Phe	
225 230 235 240	
tcg tcc cga ttg gat cct atc aag gac gat gac gag ttg gtc cgt gat	768
Ser Ser Arg Leu Asp Pro Ile Lys Asp Asp Asp Glu Leu Val Arg Asp	
245 250 255	
atc gga act aac ttg atc gtg gaa atg tgt caa aaa ttg ctc gac agt	816
Ile Gly Thr Asn Leu Ile Val Glu Met Cys Gln Lys Leu Leu Asp Ser	
260 265 270	
ggg tac gtt tct cac ttg cac atc tac acc atg aac ttg gaa aaa gcg	864
Gly Tyr Val Ser His Leu His Ile Tyr Thr Met Asn Leu Glu Lys Ala	
275 280 285	
cct ctc atg att ctg gaa aga ttg aac att cta cct acg gaa tca gag	912
Pro Leu Met Ile Leu Glu Arg Leu Asn Ile Leu Pro Thr Glu Ser Glu	
290 295 300	
ttc aat gca cat cca ttg gcc gtg ttg cca tgg aga aaa tct ttg aat	960
Phe Asn Ala His Pro Leu Ala Val Leu Pro Trp Arg Lys Ser Leu Asn	
305 310 315 320	
cca aag cgt aaa aac gag gaa gtc aga cct atc ttc tgg aag aga aga	1008
Pro Lys Arg Lys Asn Glu Glu Val Arg Pro Ile Phe Trp Lys Arg Arg	
325 330 335	
cct tac tcc tat gtc gca aga acc tct caa tgg gcc gtg gac gaa ttc	1056
Pro Tyr Ser Tyr Val Ala Arg Thr Ser Gln Trp Ala Val Asp Glu Phe	
340 345 350	
ccc aac ggt aga ttc ggt gat tcg tct tct cct gcg ttc ggt gac ttg	1104
Pro Asn Gly Arg Phe Gly Asp Ser Ser Ser Pro Ala Phe Gly Asp Leu	
355 360 365	
gat ctg tgt ggt tca gac ttg atc agg caa tca gcg aac aaa tgt ctc	1152
Asp Leu Cys Gly Ser Asp Leu Ile Arg Gln Ser Ala Asn Lys Cys Leu	
370 375 380	
gaa tta tgg tcc acc cct act tcc atc aac gac gtc gcc ttc ttg gtc	1200
Glu Leu Trp Ser Thr Pro Thr Ser Ile Asn Asp Val Ala Phe Leu Val	
385 390 395 400	
atc aac tac ttg aat gga aac ttg aag tgt tta cct tgg agt gat atc	1248
Ile Asn Tyr Leu Asn Gly Asn Leu Lys Cys Leu Pro Trp Ser Asp Ile	
405 410 415	
ccc atc aat gat gaa ata aat cca atc aaa gca cac ttg att gag ctg	1296
Pro Ile Asn Asp Glu Ile Asn Pro Ile Lys Ala His Leu Ile Glu Leu	
420 425 430	
aac cag cat tct atc atc act ata aac tct caa cct caa gtc aac ggc	1344
Asn Gln His Ser Ile Ile Thr Ile Asn Ser Gln Pro Gln Val Asn Gly	
435 440 445	
att agg tcc aat gac aaa att cat ggt tgg gga ccc aag gat ggt tac	1392
Ile Arg Ser Asn Asp Lys Ile His Gly Trp Gly Pro Lys Asp Gly Tyr	
450 455 460	

gtt tac cag aag caa tat ttg gaa ttt atg ttg ccc aag act aag ttg 1440
 Val Tyr Gln Lys Gln Tyr Leu Glu Phe Met Leu Pro Lys Thr Lys Leu
 465 470 475 480

ccc aag ttg att gac acc ttg aaa aac aat gag ttc ttg acc tac ttc 1488
 Pro Lys Leu Ile Asp Thr Leu Lys Asn Asn Glu Phe Leu Thr Tyr Phe
 485 490 495

gcc atc gac tct caa ggt gac ctg cta agt aat cat cca gac aac tcc 1536
 Ala Ile Asp Ser Gln Gly Asp Leu Leu Ser Asn His Pro Asp Asn Ser
 500 505 510

aag tcc aac gct gtg act tgg ggt att ttc ccc ggc aga gaa att ctt 1584
 Lys Ser Asn Ala Val Thr Trp Gly Ile Phe Pro Gly Arg Glu Ile Leu
 515 520 525

caa cct acc att gtc gag aaa att tcg ttc tta gcg tgg aag gag gag 1632
 Gln Pro Thr Ile Val Glu Lys Ile Ser Phe Leu Ala Trp Lys Glu Glu
 530 535 540

ttc tat cat atc ttg aat gaa tgg aaa cta aac atg aat aaa tac gat 1680
 Tyr His Ile Leu Asn Glu Trp Lys Leu Asn Met Asn Lys Tyr Asp
 545 550 555 560

aaa ccg cat agt gcc caa ttc att cag tcc ttg att gac gat tac tgc 1728
 Lys Pro His Ser Ala Gln Phe Ile Gln Ser Leu Ile Asp Asp Tyr Cys
 565 570 575

ttg gtc aat att gtt gac aat gac tac att tct cca gat gat caa atc 1776
 Leu Val Asn Ile Val Asp Asn Asp Tyr Ile Ser Pro Asp Asp Gln Ile
 580 585 590

cat tcc atc cta cta agc cta taa 1800
 His Ser Ile Leu Leu Ser Leu
 595

<210> 24

<211> 599

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 24

Met Lys Ile Thr Glu Lys Leu Glu Gln His Arg Gln Thr Ser Gly Lys
 1 5 10 15

Pro Thr Tyr Ser Phe Glu Tyr Phe Val Pro Lys Thr Thr Gln Gly Val
 20 25 30

Gln Asn Leu Tyr Asp Arg Met Asp Arg Met Tyr Glu Ala Ser Leu Pro
 35 40 45

Gln Phe Ile Asp Ile Thr Trp Asn Ala Gly Gly Gly Arg Leu Ser His
 50 55 60

Leu Ser Thr Asp Leu Val Ala Thr Ala Gln Ser Val Leu Gly Leu Glu
 65 70 75 80

Thr Cys Met His Leu Thr Cys Thr Asn Met Pro Ile Ser Met Ile Asp
 85 90 95

Asp Ala Leu Glu Asn Ala Tyr His Ser Gly Cys Gln Asn Ile Leu Ala
 100 105 110

Leu Arg Gly Asp Pro Pro Arg Asp Ala Glu Asn Trp Thr Pro Val Glu
 115 120 125
 Gly Gly Phe Gln Tyr Ala Lys Asp Leu Ile Lys Tyr Ile Lys Ser Lys
 130 135 140
 Tyr Gly Asp His Phe Ala Ile Gly Val Ala Gly Tyr Pro Glu Cys His
 145 150 155 160
 Pro Glu Leu Pro Asn Lys Asp Val Lys Leu Asp Leu Glu Tyr Leu Ser
 165 170 175
 Arg Arg Ser Thr Gly Gly Asp Phe Ile Ile Thr Gln Met Phe Tyr Asp
 180 185 190
 Val Asp Asn Leu Leu Asn Trp Cys Ser Gln Val Arg Ala Ala Gly Met
 195 200 205
 Asp Val Pro Ile Ile Pro Gly Ile Met Pro Ile Thr Thr Tyr Ala Ala
 210 215 220
 Leu Leu Arg Arg Ile Gln Trp Gly Gln Ile Ser Ile Pro Gln His Phe
 225 230 235 240
 Ser Ser Arg Leu Asp Pro Ile Lys Asp Asp Asp Glu Leu Val Arg Asp
 245 250 255
 Ile Gly Thr Asn Leu Ile Val Glu Met Cys Gln Lys Leu Leu Asp Ser
 260 265 270
 Gly Tyr Val Ser His Leu His Ile Tyr Thr Met Asn Leu Glu Lys Ala
 275 280 285
 Pro Leu Met Ile Leu Glu Arg Leu Asn Ile Leu Pro Thr Glu Ser Glu
 290 295 300
 Phe Asn Ala His Pro Leu Ala Val Leu Pro Trp Arg Lys Ser Leu Asn
 305 310 315 320
 Pro Lys Arg Lys Asn Glu Glu Val Arg Pro Ile Phe Trp Lys Arg Arg
 325 330 335
 Pro Tyr Ser Tyr Val Ala Arg Thr Ser Gln Trp Ala Val Asp Glu Phe
 340 345 350
 Pro Asn Gly Arg Phe Gly Asp Ser Ser Ser Pro Ala Phe Gly Asp Leu
 355 360 365
 Asp Leu Cys Gly Ser Asp Leu Ile Arg Gln Ser Ala Asn Lys Cys Leu
 370 375 380
 Glu Leu Trp Ser Thr Pro Thr Ser Ile Asn Asp Val Ala Phe Leu Val
 385 390 395 400
 Ile Asn Tyr Leu Asn Gly Asn Leu Lys Cys Leu Pro Trp Ser Asp Ile
 405 410 415
 Pro Ile Asn Asp Glu Ile Asn Pro Ile Lys Ala His Leu Ile Glu Leu
 420 425 430
 Asn Gln His Ser Ile Ile Thr Ile Asn Ser Gln Pro Gln Val Asn Gly
 435 440 445

Ile Arg Ser Asn Asp Lys Ile His Gly Trp Gly Pro Lys Asp Gly Tyr
450 455 460

Val Tyr Gln Lys Gln Tyr Leu Glu Phe Met Leu Pro Lys Thr Lys Leu
465 470 475 480

Pro Lys Leu Ile Asp Thr Leu Lys Asn Asn Glu Phe Leu Thr Tyr Phe
485 490 495

Ala Ile Asp Ser Gln Gly Asp Leu Leu Ser Asn His Pro Asp Asn Ser
500 505 510

Lys Ser Asn Ala Val Thr Trp Gly Ile Phe Pro Gly Arg Glu Ile Leu
515 520 525

Gln Pro Thr Ile Val Glu Lys Ile Ser Phe Leu Ala Trp Lys Glu Glu
530 535 540

Phe Tyr His Ile Leu Asn Glu Trp Lys Leu Asn Met Asn Lys Tyr Asp
545 550 555 560

Lys Pro His Ser Ala Gln Phe Ile Gln Ser Leu Ile Asp Asp Tyr Cys
565 570 575

Leu Val Asn Ile Val Asp Asn Asp Tyr Ile Ser Pro Asp Asp Gln Ile
580 585 590

His Ser Ile Leu Leu Ser Leu
595

<210> 25

<211> 897

<212> DNA

<213> Erwinia carotovora

<220>

<221> CDS

<222> (1)..(894)

<223> REO00089

<400> 25

atg agc ttt ttt cac gca aac cag cgg gaa gcg ctg aat caa agt ctg 48
Met Ser Phe Phe His Ala Asn Gln Arg Glu Ala Leu Asn Gln Ser Leu
1 5 10 15

gcg gaa ttg cag gga cga att aat gtg tca ttt gaa ttt ttc ccg cca 96
Ala Glu Leu Gln Gly Arg Ile Asn Val Ser Phe Glu Phe Phe Pro Pro
20 25 30

cgt acc agc gat atg gaa gaa acc ctg tgg agc tct atc gat cga ctg 144
Arg Thr Ser Asp Met Glu Glu Thr Leu Trp Ser Ser Ile Asp Arg Leu
35 40 45

agc agc ctg aag ccc aag ttt gtt tcc gtg act tac ggg gcg aat tct 192
Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser
50 55 60

ggc gag cgt gac cgt act cac agc att atc aaa acg att aaa gag cgt 240
Gly Glu Arg Asp Arg Thr His Ser Ile Ile Lys Thr Ile Lys Glu Arg
65 70 75 80

```

acc ggt ctg gaa gcg gca cct cac ctg acc tgc atc gat gct tca cgc 288
Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Ser Arg
      85                      90                      95

gaa cag ctg cgt gaa atc gct cag gat tac tgg gag agt ggt atc cgc 336
Glu Gln Leu Arg Glu Ile Ala Gln Asp Tyr Trp Glu Ser Gly Ile Arg
      100                      105                      110

cat att gtc gcg ctg cgc ggc gac ttg cct caa gaa ggc ggc aaa ccg 384
His Ile Val Ala Leu Arg Gly Asp Leu Pro Gln Glu Gly Gly Lys Pro
      115                      120                      125

gac atg tac gcg gcg gat ctg gtt tcc ctg ctg aaa gag gtc ggt gat 432
Asp Met Tyr Ala Ala Asp Leu Val Ser Leu Leu Lys Glu Val Gly Asp
      130                      135                      140

ttc gat att tcc gtt gcc gcc tat cct gaa gta cac cct gaa gcg aaa 480
Phe Asp Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys
      145                      150                      155                      160

gac gcg cag gct gac ctg att aac ctg aaa cac aag att gat gcc ggc 528
r Ala Gln Ala Asp Leu Ile Asn Leu Lys His Lys Ile Asp Ala Gly
      165                      170                      175

gcg aat cgc gct atc aca cag ttc ttt ttc gac gta gaa agc tat ttg 576
Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu
      180                      185                      190

cgg ttc cgt gac cgc tgc gtg gca acg ggc atc gat gta gaa att gtg 624
Arg Phe Arg Asp Arg Cys Val Ala Thr Gly Ile Asp Val Glu Ile Val
      195                      200                      205

ccg ggc att ctg cca gta tgc aac ttc aaa cag ttg cag aaa ttt gcc 672
Pro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Leu Gln Lys Phe Ala
      210                      215                      220

acg atg acc aac gtc cgt gtg ccg aac tgg atg acg acc atg ttt gac 720
Thr Met Thr Asn Val Arg Val Pro Asn Trp Met Thr Thr Met Phe Asp
      225                      230                      235                      240

ggc ctg gat aac gat cca gaa acc cgc aaa atg gtg ggg gcg tct atc 768
ly Leu Asp Asn Asp Pro Glu Thr Arg Lys Met Val Gly Ala Ser Ile
      245                      250                      255

gcc atg gat atg gtg aaa att ctc agc cgc gaa ggc gta aaa gat ttc 816
Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
      260                      265                      270

cat ttc tat acg ctg aac cgc gcg gag ctg agc tat gcg att tgc cat 864
His Phe Tyr Thr Leu Asn Arg Ala Glu Leu Ser Tyr Ala Ile Cys His
      275                      280                      285

acg ctg ggc gtc cgc cct gat gta gca cgc tga 897
Thr Leu Gly Val Arg Pro Asp Val Ala Arg
      290                      295

```

<210> 26

<211> 298

<212> PRT

<213> Erwinia carotovora

<400> 26

Met Ser Phe Phe His Ala Asn Gln Arg Glu Ala Leu Asn Gln Ser Leu
 1 5 10 15
 Ala Glu Leu Gln Gly Arg Ile Asn Val Ser Phe Glu Phe Phe Pro Pro
 20 25 30
 Arg Thr Ser Asp Met Glu Glu Thr Leu Trp Ser Ser Ile Asp Arg Leu
 35 40 45
 Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser
 50 55 60
 Gly Glu Arg Asp Arg Thr His Ser Ile Ile Lys Thr Ile Lys Glu Arg
 65 70 75 80
 Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Ser Arg
 85 90 95
 Glu Gln Leu Arg Glu Ile Ala Gln Asp Tyr Trp Glu Ser Gly Ile Arg
 100 105 110
 s Ile Val Ala Leu Arg Gly Asp Leu Pro Gln Glu Gly Gly Lys Pro
 115 120 125
 Asp Met Tyr Ala Ala Asp Leu Val Ser Leu Leu Lys Glu Val Gly Asp
 130 135 140
 Phe Asp Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys
 145 150 155 160
 Ser Ala Gln Ala Asp Leu Ile Asn Leu Lys His Lys Ile Asp Ala Gly
 165 170 175
 Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu
 180 185 190
 Arg Phe Arg Asp Arg Cys Val Ala Thr Gly Ile Asp Val Glu Ile Val
 195 200 205
 Pro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Leu Gln Lys Phe Ala
 210 215 220
 Thr Met Thr Asn Val Arg Val Pro Asn Trp Met Thr Thr Met Phe Asp
 225 230 235 240
 Gly Leu Asp Asn Asp Pro Glu Thr Arg Lys Met Val Gly Ala Ser Ile
 245 250 255
 Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
 260 265 270
 His Phe Tyr Thr Leu Asn Arg Ala Glu Leu Ser Tyr Ala Ile Cys His
 275 280 285
 Thr Leu Gly Val Arg Pro Asp Val Ala Arg
 290 295

<210> 27

<211> 888

<212> DNA

<213> Klebsiella pneumoniae

<220>
 <221> CDS
 <222> (1)..(885)
 <223> RKP07488

<400> 27

atg agc ttt ttt cac gcc aat cag cgg gaa gcc ctg aat cag agc ctg	48
Met Ser Phe Phe His Ala Asn Gln Arg Glu Ala Leu Asn Gln Ser Leu	
1 5 10 15	
gcg gaa gtc cag gcc cag att aat gtg tct ttt gaa ttc ttt ccg ccg	96
Ala Glu Val Gln Gly Gln Ile Asn Val Ser Phe Glu Phe Phe Pro Pro	
20 25 30	
cgc acc agt gaa atg gag caa acc ctg tgg aaa tcc atc gat cgc ctg	144
Arg Thr Ser Glu Met Glu Gln Thr Leu Trp Lys Ser Ile Asp Arg Leu	
35 40 45	
agc agt ctg aaa ccg aag ttt gtt tcg gta acc tat ggc gcg aac tct	192
Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser	
50 55 60	
gcg gag cgc gat cgc acc cac agc atc atc aaa ggc att aaa gag cga	240
Gly Glu Arg Asp Arg Thr His Ser Ile Ile Lys Gly Ile Lys Glu Arg	
65 70 75 80	
acc ggt ctg gaa gca gcg ccg cac ctg acc tgt atc gat gcc agc cgc	288
Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Ser Arg	
85 90 95	
gat gag ttg cgc act atc gct cag gat tac tgg aac aac ggt atc cgc	336
Asp Glu Leu Arg Thr Ile Ala Gln Asp Tyr Trp Asn Asn Gly Ile Arg	
100 105 110	
cat atc gtc gcc ctg cgc gcc gac ctg ccg ccg gcc agc ggt aaa ccg	384
His Ile Val Ala Leu Arg Gly Asp Leu Pro Pro Gly Ser Gly Lys Pro	
115 120 125	
gat atg tac gcc gcc gat ctg gtg acg ttg ctg aaa gag gta gcc gat	432
Asp Met Tyr Ala Ala Asp Leu Val Thr Leu Leu Lys Glu Val Gly Asp	
130 135 140	
ttt gat atc tct gtc gcc gcg tat ccg gaa gtg cat ccg gag gcg aaa	480
Phe Asp Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys	
145 150 155 160	
agc gcg cag gcg gat tta ctg aac ctg aag cgc aaa gta gaa gca ggg	528
Ser Ala Gln Ala Asp Leu Leu Asn Leu Lys Arg Lys Val Glu Ala Gly	
165 170 175	
gcc aac cgc gcg atc acc cag ttc ttc ttc gat gtg gaa agc tac ctg	576
Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu	
180 185 190	
cgt ttt cgc gat cgc tgc gtc tcg gca gcc atc gac gtg gaa atc att	624
Arg Phe Arg Asp Arg Cys Val Ser Ala Gly Ile Asp Val Glu Ile Ile	
195 200 205	
ccc ggt atc ctg ccg gtc tcc aac ttt aaa cag gcg aaa aag ttt gcg	672
Pro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Ala Lys Lys Phe Ala	
210 215 220	
gat atg acc aac gtc cgt atc ccg gtg tgg atg tca aaa atg ttc gaa	720

Asp Met Thr Asn Val Arg Ile Pro Val Trp Met Ser Lys Met Phe Glu
 225 230 235 240
 ggg ctg gat aac gac gcc gaa acc cgt caa ctg gtg ggg gcg aat atc 768
 Gly Leu Asp Asn Asp Ala Glu Thr Arg Gln Leu Val Gly Ala Asn Ile
 245 250 255
 gcc atg gac atg gtg aag atc tta agc cgg gaa ggg gtc aag gat ttc 816
 Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
 260 265 270
 cac ttc tac acc ctg aac cgc gcc gag atg agc tac gcc atc tgc cat 864
 His Phe Tyr Thr Leu Asn Arg Ala Glu Met Ser Tyr Ala Ile Cys His
 275 280 285
 acg ctg ggc gta cgc ccg gcc tga 888
 Thr Leu Gly Val Arg Pro Ala
 290 295

<210> 28
 <211> 295
 <212> PRT

<213> *Klebsiella pneumoniae*

<400> 28

Met Ser Phe Phe His Ala Asn Gln Arg Glu Ala Leu Asn Gln Ser Leu
 1 5 10 15
 Ala Glu Val Gln Gly Gln Ile Asn Val Ser Phe Glu Phe Phe Pro Pro
 20 25 30
 Arg Thr Ser Glu Met Glu Gln Thr Leu Trp Lys Ser Ile Asp Arg Leu
 35 40 45
 Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser
 50 55 60
 Gly Glu Arg Asp Arg Thr His Ser Ile Ile Lys Gly Ile Lys Glu Arg
 65 70 75 80
 Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Ser Arg
 85 90 95
 Asp Glu Leu Arg Thr Ile Ala Gln Asp Tyr Trp Asn Asn Gly Ile Arg
 100 105 110
 His Ile Val Ala Leu Arg Gly Asp Leu Pro Pro Gly Ser Gly Lys Pro
 115 120 125
 Asp Met Tyr Ala Ala Asp Leu Val Thr Leu Leu Lys Glu Val Gly Asp
 130 135 140
 Phe Asp Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys
 145 150 155 160
 Ser Ala Gln Ala Asp Leu Leu Asn Leu Lys Arg Lys Val Glu Ala Gly
 165 170 175
 Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu
 180 185 190
 Arg Phe Arg Asp Arg Cys Val Ser Ala Gly Ile Asp Val Glu Ile Ile

195

200

205

Pro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Ala Lys Lys Phe Ala
 210 215 220

Asp Met Thr Asn Val Arg Ile Pro Val Trp Met Ser Lys Met Phe Glu
 225 230 235 240

Gly Leu Asp Asn Asp Ala Glu Thr Arg Gln Leu Val Gly Ala Asn Ile
 245 250 255

Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
 260 265 270

His Phe Tyr Thr Leu Asn Arg Ala Glu Met Ser Tyr Ala Ile Cys His
 275 280 285

Thr Leu Gly Val Arg Pro Ala
 290 295

<210> 29

<211> 891

<212> DNA

<213> Salmonella typhi

<220>

<221> CDS

<222> (1)..(888)

<223> RTY02485

<400> 29

atg agc ttt ttt cac gcc aac cag cgg gaa gcc ctg aat cag agc ctg 48
 Met Ser Phe Phe His Ala Asn Gln Arg Glu Ala Leu Asn Gln Ser Leu
 1 5 10 15

gcg gaa gta cag ggt cag att aac gtt tcg ttt gaa ttt ttc ccg ccg 96
 Ala Glu Val Gln Gly Gln Ile Asn Val Ser Phe Glu Phe Phe Pro Pro
 20 25 30

cgc acc agt gaa atg gag caa acc ctg tgg aac tcc atc gat cgc ctg 144
 Arg Thr Ser Glu Met Glu Gln Thr Leu Trp Asn Ser Ile Asp Arg Leu
 35 40 45

agc agc ctg aaa ccg aag ttt gtt tcg gta acg tat ggc gcc aac tcc 192
 Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser
 50 55 60

ggg gaa cgt gac cgc act cat agt gtt att aaa ggc att aaa gag cgt 240
 Gly Glu Arg Asp Arg Thr His Ser Val Ile Lys Gly Ile Lys Glu Arg
 65 70 75 80

act ggg ctt gag gcc gcg ccg cac ctt acc tgt att gac gcc acg cgc 288
 Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Thr Arg
 85 90 95

gat gaa ctg cgc acc atc gcc cgc gac tac tgg aat aac ggt atc cgc 336
 Asp Glu Leu Arg Thr Ile Ala Arg Asp Tyr Trp Asn Asn Gly Ile Arg
 100 105 110

cac att gtt gct ttg cgc ggc gat ttg ccg ccg ggc agc ggt aag ccg 384
 His Ile Val Ala Leu Arg Gly Asp Leu Pro Pro Gly Ser Gly Lys Pro
 115 120 125

gag atg tac gcc gcc gat ctg gtt ggt ttg ctc aaa gag gtg gtc gat 432
 Glu Met Tyr Ala Ala Asp Leu Val Gly Leu Leu Lys Glu Val Val Asp
 130 135 140

ttc gat att tca gta gcg gcc tat ccg gag gta cat ccg gaa gcg aaa 480
 Phe Asp Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys
 145 150 155 160

agc gcg cag gcc gat ctg ctt aat ctg aag cgt aaa gtg gat gct ggc 528
 Ser Ala Gln Ala Asp Leu Leu Asn Leu Lys Arg Lys Val Asp Ala Gly
 165 170 175

gct aac cgc gcg ata acc caa ttt ttc ttc gat gtg gaa agc tat ctg 576
 Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu
 180 185 190

cgt ttt cgc gac cgc tgt gtt tcc gcc ggt atc gac gta gaa att att 624
 Arg Phe Arg Asp Arg Cys Val Ser Ala Gly Ile Asp Val Glu Ile Ile
 195 200 205

c ggc att tta ccg gtg tct aac ttt aaa cag gcg aaa aaa ttt gcc 672
 ro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Ala Lys Lys Phe Ala
 210 215 220

gat atg acc aat gtc cgc att ccg tcc tgg atg tgc ctg atg ttt gag 720
 Asp Met Thr Asn Val Arg Ile Pro Ser Trp Met Ser Leu Met Phe Glu
 225 230 235 240

ggg ctg gat gat gac gca gaa acc cgc aag ctg gtg ggc gct aac att 768
 Gly Leu Asp Asp Asp Ala Glu Thr Arg Lys Leu Val Gly Ala Asn Ile
 245 250 255

gcg atg gac atg gtg aaa att tta agc cgc gaa gga gtg aag gat ttc 816
 Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
 260 265 270

cac ttc tac acg ttg aat cgt gcg gaa atg agt tat gcc att tgc cac 864
 His Phe Tyr Thr Leu Asn Arg Ala Glu Met Ser Tyr Ala Ile Cys His
 275 280 285

cg ctg ggc gta aga ccg ggt tta taa 891
 hr Leu Gly Val Arg Pro Gly Leu
 290 295

<210> 30

<211> 296

<212> PRT

<213> Salmonella typhi

<400> 30

Met Ser Phe Phe His Ala Asn Gln Arg Glu Ala Leu Asn Gln Ser Leu
1 5 10 15

Ala Glu Val Gln Gly Gln Ile Asn Val Ser Phe Glu Phe Phe Pro Pro
20 25 30

Arg Thr Ser Glu Met Glu Gln Thr Leu Trp Asn Ser Ile Asp Arg Leu
35 40 45

Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser
50 55 60

Gly Glu Arg Asp Arg Thr His Ser Val Ile Lys Gly Ile Lys Glu Arg
 65 70 75 80
 Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Thr Arg
 85 90 95
 Asp Glu Leu Arg Thr Ile Ala Arg Asp Tyr Trp Asn Asn Gly Ile Arg
 100 105 110
 His Ile Val Ala Leu Arg Gly Asp Leu Pro Pro Gly Ser Gly Lys Pro
 115 120 125
 Glu Met Tyr Ala Ala Asp Leu Val Gly Leu Leu Lys Glu Val Val Asp
 130 135 140
 Phe Asp Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys
 145 150 155 160
 Ser Ala Gln Ala Asp Leu Leu Asn Leu Lys Arg Lys Val Asp Ala Gly
 165 170 175
 Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu
 180 185 190
 Arg Phe Arg Asp Arg Cys Val Ser Ala Gly Ile Asp Val Glu Ile Ile
 195 200 205
 Pro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Ala Lys Lys Phe Ala
 210 215 220
 Asp Met Thr Asn Val Arg Ile Pro Ser Trp Met Ser Leu Met Phe Glu
 225 230 235 240
 Gly Leu Asp Asp Asp Ala Glu Thr Arg Lys Leu Val Gly Ala Asn Ile
 245 250 255
 Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
 260 265 270
 His Phe Tyr Thr Leu Asn Arg Ala Glu Met Ser Tyr Ala Ile Cys His
 275 280 285
 Thr Leu Gly Val Arg Pro Gly Leu
 290 295

<210> 31
 <211> 891
 <212> DNA
 <213> Salmonella typhimurium

<220>
 <221> CDS
 <222> (1)..(888)
 <223> RSY00593

<400> 31
 atg agc ttt ttt cac gcc aac cag cgg gaa gcc ctg aat cag agc ctg 48
 Met Ser Phe Phe His Ala Asn Gln Arg Glu Ala Leu Asn Gln Ser Leu
 1 5 10 15
 gcg gaa gta cag ggt cag att aac gtt tcg ttt gaa ttt ttc ccg ccg 96

Ala	Glu	Val	Gln	Gly	Gln	Ile	Asn	Val	Ser	Phe	Glu	Phe	Phe	Pro	Pro		
			20					25						30			
cgc	acc	agt	gaa	atg	gag	caa	acc	ctg	tgg	aac	tcc	atc	gat	cgc	ctg	144	
Arg	Thr	Ser	Glu	Met	Glu	Gln	Thr	Leu	Trp	Asn	Ser	Ile	Asp	Arg	Leu		
		35					40				45						
agc	agt	ctg	aaa	ccg	aag	ttt	gtt	tcg	gta	acg	tat	ggc	gcc	aac	tcc	192	
Ser	Ser	Leu	Lys	Pro	Lys	Phe	Val	Ser	Val	Thr	Tyr	Gly	Ala	Asn	Ser		
	50					55				60							
ggg	gaa	cgc	gac	cgc	acc	cat	agc	gtt	att	aaa	ggc	atc	aaa	gag	cgt	240	
Gly	Glu	Arg	Asp	Arg	Thr	His	Ser	Val	Ile	Lys	Gly	Ile	Lys	Glu	Arg		
65					70				75						80		
act	ggg	ctt	gag	gcc	gcg	ccg	cac	ctt	acc	tgt	att	gac	gcc	acg	cgc	288	
Thr	Gly	Leu	Glu	Ala	Ala	Pro	His	Leu	Thr	Cys	Ile	Asp	Ala	Thr	Arg		
			85					90						95			
gat	gaa	ctg	cgc	acc	atc	gcc	cgc	gac	tac	tgg	aac	aac	ggg	atc	cgc	336	
asp	Glu	Leu	Arg	Thr	Ile	Ala	Arg	Asp	Tyr	Trp	Asn	Asn	Gly	Ile	Arg		
		100						105					110				
cac	att	gtc	gct	ttg	cgc	ggc	gat	ttg	ccg	ccg	ggc	agc	ggg	aag	ccg	384	
His	Ile	Val	Ala	Leu	Arg	Gly	Asp	Leu	Pro	Pro	Gly	Ser	Gly	Lys	Pro		
		115					120					125					
gag	atg	tac	gcc	gcc	gat	ctg	gtt	ggg	ttg	ctc	aaa	gag	gtg	gcc	gat	432	
Glu	Met	Tyr	Ala	Ala	Asp	Leu	Val	Gly	Leu	Leu	Lys	Glu	Val	Ala	Asp		
	130					135					140						
ttc	gat	att	tca	gta	gcg	gcc	tat	ccg	gag	gta	cat	ccg	gaa	gcg	aaa	480	
Phe	Asp	Ile	Ser	Val	Ala	Ala	Tyr	Pro	Glu	Val	His	Pro	Glu	Ala	Lys		
145					150				155						160		
agc	gcg	cag	gcc	gat	ctg	ctt	aac	ctg	aag	cgt	aaa	gtg	gat	gct	ggc	528	
Ser	Ala	Gln	Ala	Asp	Leu	Leu	Asn	Leu	Lys	Arg	Lys	Val	Asp	Ala	Gly		
			165					170					175				
gct	aac	cgc	gcg	ata	acc	caa	ttt	ttc	ttc	gat	gtg	gaa	agc	tac	ctg	576	
Ala	Asn	Arg	Ala	Ile	Thr	Gln	Phe	Phe	Phe	Asp	Val	Glu	Ser	Tyr	Leu		
		180						185					190				
cgt	ttt	cgc	gac	cgc	tgt	gtt	tct	gcc	ggg	atc	gac	gta	gaa	att	att	624	
Arg	Phe	Arg	Asp	Arg	Cys	Val	Ser	Ala	Gly	Ile	Asp	Val	Glu	Ile	Ile		
		195				200						205					
ccc	ggc	att	tta	ccg	gtg	tct	aac	ttt	aaa	cag	gca	aaa	aaa	ttt	gcc	672	
Pro	Gly	Ile	Leu	Pro	Val	Ser	Asn	Phe	Lys	Gln	Ala	Lys	Lys	Phe	Ala		
	210					215					220						
gat	atg	acc	aac	gtc	cgc	att	ccg	tcc	tgg	atg	tca	ctg	atg	ttt	gag	720	
Asp	Met	Thr	Asn	Val	Arg	Ile	Pro	Ser	Trp	Met	Ser	Leu	Met	Phe	Glu		
	225				230				235					240			
ggg	ctg	gat	aac	gac	gca	gaa	acc	cgc	aag	ctg	gtg	ggc	gct	aac	att	768	
Gly	Leu	Asp	Asn	Asp	Ala	Glu	Thr	Arg	Lys	Leu	Val	Gly	Ala	Asn	Ile		
			245					250					255				
gcg	atg	gac	atg	gtg	aaa	att	tta	agc	cgt	gaa	gga	gtg	aag	gat	ttc	816	
Ala	Met	Asp	Met	Val	Lys	Ile	Leu	Ser	Arg	Glu	Gly	Val	Lys	Asp	Phe		
		260					265						270				

cac ttc tac acg ttg aat cgt gcg gaa atg agt tat gcc att tgc cac 864
 His Phe Tyr Thr Leu Asn Arg Ala Glu Met Ser Tyr Ala Ile Cys His
 275 280 285

acg ctg ggc gta aga ccg ggt tta taa 891
 Thr Leu Gly Val Arg Pro Gly Leu
 290 295

<210> 32

<211> 296

<212> PRT

<213> Salmonella typhimurium

<400> 32

Met Ser Phe Phe His Ala Asn Gln Arg Glu Ala Leu Asn Gln Ser Leu
 1 5 10 15

Ala Glu Val Gln Gly Gln Ile Asn Val Ser Phe Glu Phe Phe Pro Pro
 20 25 30

G Thr Ser Glu Met Glu Gln Thr Leu Trp Asn Ser Ile Asp Arg Leu
 35 40 45

Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser
 50 55 60

Gly Glu Arg Asp Arg Thr His Ser Val Ile Lys Gly Ile Lys Glu Arg
 65 70 75 80

Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Thr Arg
 85 90 95

Asp Glu Leu Arg Thr Ile Ala Arg Asp Tyr Trp Asn Asn Gly Ile Arg
 100 105 110

His Ile Val Ala Leu Arg Gly Asp Leu Pro Pro Gly Ser Gly Lys Pro
 115 120 125

Glu Met Tyr Ala Ala Asp Leu Val Gly Leu Leu Lys Glu Val Ala Asp
 130 135 140

he Asp Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys
 145 150 155 160

Ser Ala Gln Ala Asp Leu Leu Asn Leu Lys Arg Lys Val Asp Ala Gly
 165 170 175

Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu
 180 185 190

Arg Phe Arg Asp Arg Cys Val Ser Ala Gly Ile Asp Val Glu Ile Ile
 195 200 205

Pro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Ala Lys Lys Phe Ala
 210 215 220

Asp Met Thr Asn Val Arg Ile Pro Ser Trp Met Ser Leu Met Phe Glu
 225 230 235 240

Gly Leu Asp Asn Asp Ala Glu Thr Arg Lys Leu Val Gly Ala Asn Ile
 245 250 255

Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
 260 265 270

His Phe Tyr Thr Leu Asn Arg Ala Glu Met Ser Tyr Ala Ile Cys His
 275 280 285

Thr Leu Gly Val Arg Pro Gly Leu
 290 295

<210> 33
 <211> 891
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (1) .. (888)
 <223> REC03839

400> 33
 g agc ttt ttt cac gcc agc cag cgg gat gcc ctg aat cag agc ctg 48
 et Ser Phe Phe His Ala Ser Gln Arg Asp Ala Leu Asn Gln Ser Leu
 1 5 10 15

gca gaa gtc cag ggg cag att aac gtt tgc ttc gag ttt ttc ccg ccg 96
 Ala Glu Val Gln Gly Gln Ile Asn Val Ser Phe Glu Phe Phe Pro Pro
 20 25 30

cgt acc agt gaa atg gag cag acc ctg tgg aac tcc atc gat cgc ctt 144
 Arg Thr Ser Glu Met Glu Gln Thr Leu Trp Asn Ser Ile Asp Arg Leu
 35 40 45

agc agc ctg aaa ccg aag ttt gta tgc gtg acc tat ggc gcg aac tcc 192
 Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser
 50 55 60

ggc gag cgc gac cgt acg cac agc att att aaa ggc att aaa gat cgc 240
 Gly Glu Arg Asp Arg Thr His Ser Ile Ile Lys Gly Ile Lys Asp Arg
 65 70 75 80

ct ggt ctg gaa gcg gca ccg cat ctt act tgc att gat gcg acg ccc 288
 Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Thr Pro
 85 90 95

gac gag ctg cgc acc att gca cgc gac tac tgg aat aac ggt att cgt 336
 Asp Glu Leu Arg Thr Ile Ala Arg Asp Tyr Trp Asn Asn Gly Ile Arg
 100 105 110

cat atc gtg gcg ctg cgt ggc gat ctg ccg ccg gga agt ggt aag cca 384
 His Ile Val Ala Leu Arg Gly Asp Leu Pro Pro Gly Ser Gly Lys Pro
 115 120 125

gaa atg tat gct tct gac ctg gtg acg ctg tta aaa gaa gtg gca gat 432
 Glu Met Tyr Ala Ser Asp Leu Val Thr Leu Leu Lys Glu Val Ala Asp
 130 135 140

ttc gat atc tcc gtg gcg gcg tat ccg gaa gtt cac ccg gaa gca aaa 480
 Phe Asp Ile Ser Val Ala Tyr Pro Glu Val His Pro Glu Ala Lys
 145 150 155 160

agc gct cag gcg gat ttg ctt aat ctg aaa cgc aaa gtg gat gcc gga 528

Ser Ala Gln Ala Asp Leu Leu Asn Leu Lys Arg Lys Val Asp Ala Gly
165 170 175

gcc aac cgc gcg att act cag ttc ttc ttc gat gtc gaa agc tac ctg 576
Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu
180 185 190

cgt ttt cgt gac cgc tgt gta tgc gcg ggc att gat gtg gaa att att 624
Arg Phe Arg Asp Arg Cys Val Ser Ala Gly Ile Asp Val Glu Ile Ile
195 200 205

ccg gga att ttg ccg gta tct aac ttt aaa cag gcg aag aaa ttt gcc 672
Pro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Ala Lys Lys Phe Ala
210 215 220

gat atg acc aac gtg cgt att ccg gcg tgg atg gcg caa atg ttc gac 720
Asp Met Thr Asn Val Arg Ile Pro Ala Trp Met Ala Gln Met Phe Asp
225 230 235 240

ggc ctg gat gat gat gcc gaa acc cgc aaa ctg gtt ggc gcg aat att 768
Gly Leu Asp Asp Asp Ala Glu Thr Arg Lys Leu Val Gly Ala Asn Ile
245 250 255

gcc atg gat atg gtg aag att tta agc cgt gaa gga gtg aaa gat ttc 816
Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
260 265 270

cac ttc tat acg ctt aac cgt gct gaa atg agt tac gcg att tgc cat 864
His Phe Tyr Thr Leu Asn Arg Ala Glu Met Ser Tyr Ala Ile Cys His
275 280 285

acg ctg ggg gtt cga cct ggt tta taa 891
Thr Leu Gly Val Arg Pro Gly Leu
290 295

<210> 34

<211> 296

<212> PRT

<213> Escherichia coli

<400> 34

Met Ser Phe Phe His Ala Ser Gln Arg Asp Ala Leu Asn Gln Ser Leu
1 5 10 15

Ala Glu Val Gln Gly Gln Ile Asn Val Ser Phe Glu Phe Phe Pro Pro
20 25 30

Arg Thr Ser Glu Met Glu Gln Thr Leu Trp Asn Ser Ile Asp Arg Leu
35 40 45

Ser Ser Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser
50 55 60

Gly Glu Arg Asp Arg Thr His Ser Ile Ile Lys Gly Ile Lys Asp Arg
65 70 75 80

Thr Gly Leu Glu Ala Ala Pro His Leu Thr Cys Ile Asp Ala Thr Pro
85 90 95

Asp Glu Leu Arg Thr Ile Ala Arg Asp Tyr Trp Asn Asn Gly Ile Arg
100 105 110

His Ile Val Ala Leu Arg Gly Asp Leu Pro Pro Gly Ser Gly Lys Pro
 115 120 125
 Glu Met Tyr Ala Ser Asp Leu Val Thr Leu Leu Lys Glu Val Ala Asp
 130 135 140
 Phe Asp Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys
 145 150 155 160
 Ser Ala Gln Ala Asp Leu Leu Asn Leu Lys Arg Lys Val Asp Ala Gly
 165 170 175
 Ala Asn Arg Ala Ile Thr Gln Phe Phe Phe Asp Val Glu Ser Tyr Leu
 180 185 190
 Arg Phe Arg Asp Arg Cys Val Ser Ala Gly Ile Asp Val Glu Ile Ile
 195 200 205
 Pro Gly Ile Leu Pro Val Ser Asn Phe Lys Gln Ala Lys Lys Phe Ala
 210 215 220
 p Met Thr Asn Val Arg Ile Pro Ala Trp Met Ala Gln Met Phe Asp
 225 230 235 240
 Gly Leu Asp Asp Asp Ala Glu Thr Arg Lys Leu Val Gly Ala Asn Ile
 245 250 255
 Ala Met Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Lys Asp Phe
 260 265 270
 His Phe Tyr Thr Leu Asn Arg Ala Glu Met Ser Tyr Ala Ile Cys His
 275 280 285
 Thr Leu Gly Val Arg Pro Gly Leu
 290 295

<210> 35

<211> 915

<212> DNA

<213> Vibrio cholerae

<220>

<221> CDS

<222> (1)..(912)

<223> RVC06433

<400> 35

gtg aca ctc ggt cac agg gag tac aag atg gga tac aca cac gct agc 48
 Val Thr Leu Gly His Arg Glu Tyr Lys Met Gly Tyr Thr His Ala Ser
 1 5 10 15

cat atc gat gca ttg aac caa aac att gcg gag ctt tcc gac atc aat 96
 His Ile Asp Ala Leu Asn Gln Asn Ile Ala Glu Leu Ser Asp Ile Asn
 20 25 30

gtt tcg ttt gag ttt ttt cca ccc agc tca cca caa atg gaa gaa acg 144
 Val Ser Phe Glu Phe Phe Pro Pro Ser Ser Pro Gln Met Glu Glu Thr
 35 40 45

ctt tgg gga tcg gta cac cgt ctg aaa aca ctc caa ccg aaa ttt gtt 192
 Leu Trp Gly Ser Val His Arg Leu Lys Thr Leu Gln Pro Lys Phe Val
 50 55 60

tcg gtc act tat ggt gca aac tct ggt gag cgt gac cgt act cac tcg	240
Ser Val Thr Tyr Gly Ala Asn Ser Gly Glu Arg Asp Arg Thr His Ser	
65 70 75 80	
atc att aaa gcg atc aaa gat caa acc ggt tta att gcc gcg cca cac	288
Ile Ile Lys Ala Ile Lys Asp Gln Thr Gly Leu Ile Ala Ala Pro His	
85 90 95	
ctg act tgt atc gat gcc act cgt gat gaa ctg atc cag atc gcc gat	336
Leu Thr Cys Ile Asp Ala Thr Arg Asp Glu Leu Ile Gln Ile Ala Asp	
100 105 110	
gac tac tgg cat aac ggc atc cag aat att gtg gcg ctg cgt ggg gat	384
Asp Tyr Trp His Asn Gly Ile Gln Asn Ile Val Ala Leu Arg Gly Asp	
115 120 125	
atc ccg gct ggc ggt ggt aag cca gag atg tac gcc tcc gat cta gtg	432
Ile Pro Ala Gly Gly Gly Lys Pro Glu Met Tyr Ala Ser Asp Leu Val	
130 135 140	
g ctg ctc aaa tca cgc cac gat ttt gat att tcc gtg gcc gcc ttc	480
Leu Leu Lys Ser Arg His Asp Phe Asp Ile Ser Val Ala Ala Phe	
145 150 155 160	
cct gaa gtg cac cct gaa gcc aaa agc gcg caa gcg gac ctg ctc aat	528
Pro Glu Val His Pro Glu Ala Lys Ser Ala Gln Ala Asp Leu Leu Asn	
165 170 175	
tta aaa cgt aaa gtc gat gca ggt gcg aat cgt gcc atc acg cag ttt	576
Leu Lys Arg Lys Val Asp Ala Gly Ala Asn Arg Ala Ile Thr Gln Phe	
180 185 190	
ttc ttt gat gta gaa agc tac ctg cgt ttt cgc gat cgc tgt gtg gcc	624
Phe Phe Asp Val Glu Ser Tyr Leu Arg Phe Arg Asp Arg Cys Val Ala	
195 200 205	
gct ggg att gac gta gaa atc gtg cct ggc att ctg ccg gtt tct aac	672
Ala Gly Ile Asp Val Glu Ile Val Pro Gly Ile Leu Pro Val Ser Asn	
210 215 220	
ttt aaa caa gcg tcg cgc ttc gct gcg caa aac aac gtc aaa gtt ccg	720
Leu Lys Gln Ala Ser Arg Phe Ala Ala Gln Asn Asn Val Lys Val Pro	
225 230 235 240	
aat tgg atg gtg aag cag ttt gaa gga tta gaa gac gat cca gtg act	768
Asn Trp Met Val Lys Gln Phe Glu Gly Leu Glu Asp Asp Pro Val Thr	
245 250 255	
cgc cag ttg gta ggt gca agc caa gcc att gat atg gtg cgc gtg ctg	816
Arg Gln Leu Val Gly Ala Ser Gln Ala Ile Asp Met Val Arg Val Leu	
260 265 270	
tgc cgt gaa ggg gtg aag gat ttc cac ttc tac acc cta aat cgt gcc	864
Cys Arg Glu Gly Val Lys Asp Phe His Phe Tyr Thr Leu Asn Arg Ala	
275 280 285	
gaa atg act tac gcg tta tgc cac acc tta ggc gtt cgc cca caa gct	912
Glu Met Thr Tyr Ala Leu Cys His Thr Leu Gly Val Arg Pro Gln Ala	
290 295 300	
taa	915

<210> 36
 <211> 304
 <212> PRT
 <213> *Vibrio cholerae*

<400> 36
 Val Thr Leu Gly His Arg Glu Tyr Lys Met Gly Tyr Thr His Ala Ser
 1 5 10 15
 His Ile Asp Ala Leu Asn Gln Asn Ile Ala Glu Leu Ser Asp Ile Asn
 20 25 30
 Val Ser Phe Glu Phe Phe Pro Pro Ser Ser Pro Gln Met Glu Glu Thr
 35 40 45
 Leu Trp Gly Ser Val His Arg Leu Lys Thr Leu Gln Pro Lys Phe Val
 50 55 60
 Ser Val Thr Tyr Gly Ala Asn Ser Gly Glu Arg Asp Arg Thr His Ser
 65 70 75 80
 Leu Ile Lys Ala Ile Lys Asp Gln Thr Gly Leu Ile Ala Ala Pro His
 85 90 95
 Leu Thr Cys Ile Asp Ala Thr Arg Asp Glu Leu Ile Gln Ile Ala Asp
 100 105 110
 Asp Tyr Trp His Asn Gly Ile Gln Asn Ile Val Ala Leu Arg Gly Asp
 115 120 125
 Ile Pro Ala Gly Gly Gly Lys Pro Glu Met Tyr Ala Ser Asp Leu Val
 130 135 140
 Thr Leu Leu Lys Ser Arg His Asp Phe Asp Ile Ser Val Ala Ala Phe
 145 150 155 160
 Pro Glu Val His Pro Glu Ala Lys Ser Ala Gln Ala Asp Leu Leu Asn
 165 170 175
 Leu Lys Arg Lys Val Asp Ala Gly Ala Asn Arg Ala Ile Thr Gln Phe
 180 185 190
 Phe Phe Asp Val Glu Ser Tyr Leu Arg Phe Arg Asp Arg Cys Val Ala
 195 200 205
 Ala Gly Ile Asp Val Glu Ile Val Pro Gly Ile Leu Pro Val Ser Asn
 210 215 220
 Phe Lys Gln Ala Ser Arg Phe Ala Ala Gln Asn Asn Val Lys Val Pro
 225 230 235 240
 Asn Trp Met Val Lys Gln Phe Glu Gly Leu Glu Asp Asp Pro Val Thr
 245 250 255
 Arg Gln Leu Val Gly Ala Ser Gln Ala Ile Asp Met Val Arg Val Leu
 260 265 270
 Cys Arg Glu Gly Val Lys Asp Phe His Phe Tyr Thr Leu Asn Arg Ala
 275 280 285
 Glu Met Thr Tyr Ala Leu Cys His Thr Leu Gly Val Arg Pro Gln Ala
 290 295 300

<210> 37
 <211> 879
 <212> DNA
 <213> Haemophilus influenzae

<220>
 <221> CDS
 <222> (1)..(876)
 <223> RH106620

```

<400> 37
atg agc tac gcg aaa gaa att gat aca tta aat caa cat att gca gat   48
Met Ser Tyr Ala Lys Glu Ile Asp Thr Leu Asn Gln His Ile Ala Asp
   1             5             10             15

ttt aat aaa aaa att aat gtc tcc ttt gaa ttt ttt cca cct aaa aac   96
Leu Asn Lys Lys Ile Asn Val Ser Phe Glu Phe Phe Pro Pro Lys Asn
           20             25             30

gaa aaa atg gaa acc ctt cta tgg gat tca att cat cgt tta aaa gta   144
Glu Lys Met Glu Thr Leu Leu Trp Asp Ser Ile His Arg Leu Lys Val
           35             40             45

tta aag cct aaa ttt gtg tca gtc act tac ggt gca aat tcg gga gaa   192
Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser Gly Glu
           50             55             60

cgt gac cgc act cac ggc att gtg aaa gcc att aaa caa gaa act ggc   240
Arg Asp Arg Thr His Gly Ile Val Lys Ala Ile Lys Gln Glu Thr Gly
           65             70             75             80

tta gaa gcc gca cca cat tta act gga att gat gcc aca cct gaa gaa   288
Leu Glu Ala Ala Pro His Leu Thr Gly Ile Asp Ala Thr Pro Glu Glu
           85             90             95

tta aaa caa att gcg aga gat tat tgg gat agt ggt att cgc cgt att   336
Leu Lys Gln Ile Ala Arg Asp Tyr Trp Asp Ser Gly Ile Arg Arg Ile
           100            105            110

gtt gcg tta cgc ggt gac gaa cct aaa ggt tac gcg aaa aaa cca ttt   384
Val Ala Leu Arg Gly Asp Glu Pro Lys Gly Tyr Ala Lys Lys Pro Phe
           115            120            125

tat gcg tca gat ctt gtg gaa tta ctc cgt tct gtc gct gat ttt gat   432
Tyr Ala Ser Asp Leu Val Glu Leu Leu Arg Ser Val Ala Asp Phe Asp
           130            135            140

att tct gta gcc gct tat ccc gaa gtt cat cca gaa gca aaa tcc gca   480
Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys Ser Ala
           145            150            155            160

caa gca gac tta att aat tta aaa cgt aaa att gat gca ggt gca aac   528
Gln Ala Asp Leu Ile Asn Leu Lys Arg Lys Ile Asp Ala Gly Ala Asn
           165            170            175

cac gtc att aca caa ttt ttc ttt gat att gaa aac tac cta cgt ttt   576
His Val Ile Thr Gln Phe Phe Phe Asp Ile Glu Asn Tyr Leu Arg Phe
           180            185            190

```


Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys Ser Ala
 145 150 155 160
 Gln Ala Asp Leu Ile Asn Leu Lys Arg Lys Ile Asp Ala Gly Ala Asn
 165 170 175
 His Val Ile Thr Gln Phe Phe Phe Asp Ile Glu Asn Tyr Leu Arg Phe
 180 185 190
 Arg Asp Arg Cys Ala Ser Ile Gly Ile Asp Thr Glu Ile Val Pro Gly
 195 200 205
 Ile Leu Pro Val Thr Asn Phe Lys Gln Leu Gln Lys Met Ala Ser Phe
 210 215 220
 Thr Asn Val Lys Ile Pro Ala Trp Leu Val Lys Ala Tyr Asp Gly Leu
 225 230 235 240
 Asp Asn Asp Pro Thr Thr Arg Asn Leu Val Ala Ala Ser Val Ala Met
 245 250 255
 Asp Met Val Lys Ile Leu Ser Arg Glu Gly Val Asn Asp Phe His Phe
 260 265 270
 Tyr Thr Leu Asn Arg Ser Glu Leu Thr Tyr Ala Ile Cys His Met Leu
 275 280 285
 Gly Val Arg Pro
 290

<210> 39
 <211> 945
 <212> DNA
 <213> *Caulobacter crescentus*

<220>
 <221> CDS
 <222> (1)..(942)
 <223> RCO02274

<400> 39
 atg acc ctt ccg ccc acc cgc cgc gtg atc ggt ccc gtc gcc cga gcc 48
 Met Thr Leu Pro Pro Thr Arg Arg Val Ile Gly Pro Val Ala Arg Ala
 1 5 10 15
 ggc gag cgg acc ggc cgt ccg cgc gtg tcg ttc gag ttc ttc ccg ccc 96
 Gly Glu Arg Thr Gly Arg Pro Arg Val Ser Phe Glu Phe Phe Pro Pro
 20 25 30
 aag act ccg cag atg gaa gag agc ctg tgg cag gcg atc aca cgc ctg 144
 Lys Thr Pro Gln Met Glu Glu Ser Leu Trp Gln Ala Ile Thr Arg Leu
 35 40 45
 gcg ccg ctg gat ccg gcc ttc gtc tcg gtg acc tat ggc gcg ggc ggc 192
 Ala Pro Leu Asp Pro Ala Phe Val Ser Val Thr Tyr Gly Ala Gly Gly
 50 55 60
 tcc acc cgc gag cgc acc cac cgc acc gtc aag cgg atc ctg gac gag 240
 Ser Thr Arg Glu Arg Thr His Arg Thr Val Lys Arg Ile Leu Asp Glu
 65 70 75 80

acc agc ctc aag ccc gcc gcg cac ctg acc tgc gtc ggc gcc agt cgc	288
Thr Ser Leu Lys Pro Ala Ala His Leu Thr Cys Val Gly Ala Ser Arg	
85 90 95	
gaa gag gtc gat gag gtc att cgc gag tac tgg gag acc ggg gtc cgt	336
Glu Glu Val Asp Glu Val Ile Arg Glu Tyr Trp Glu Thr Gly Val Arg	
100 105 110	
cac atc gtt tgc ctg cgg ggc gat ccg ccg ccc ggc gag ggc ggc atc	384
His Ile Val Ser Leu Arg Gly Asp Pro Pro Pro Gly Glu Gly Gly Ile	
115 120 125	
ggc ggg gtc tat gtg ccg cgc gcc gac ggc tac gcc aac gcc aca gag	432
Gly Gly Val Tyr Val Pro Arg Ala Asp Gly Tyr Ala Asn Ala Thr Glu	
130 135 140	
ttg acc aag gcc gtg cgc gcg atc gcg ccg ttc gag gtg ctg gtc ggg	480
Leu Thr Lys Ala Val Arg Ala Ile Ala Pro Phe Glu Val Leu Val Gly	
145 150 155 160	
ctc tat ccc gag aag cat ccc gag agc ccc tgc ttg gag cac gac atc	528
Al Tyr Pro Glu Lys His Pro Glu Ser Pro Ser Leu Glu His Asp Ile	
165 170 175	
gac gtc ttg aag cag aag gtc gac gcc ggc gcg acg ctg ggg atc agc	576
Asp Val Leu Lys Gln Lys Val Asp Ala Gly Ala Thr Leu Gly Ile Ser	
180 185 190	
cag ttc ttc ttc gac ctc gac gcc ttc ctg cgc ttc gtc gac aag gtg	624
Gln Phe Phe Phe Asp Leu Asp Ala Phe Leu Arg Phe Val Asp Lys Val	
195 200 205	
cgc gcg gcg ggc atc acc att ccg atc gtg ccg ggg atc atg ccg gtg	672
Arg Ala Ala Gly Ile Thr Ile Pro Ile Val Pro Gly Ile Met Pro Val	
210 215 220	
acc aat ttc gcg ggc ttg aag aag atg gcc gcc gcc tgc cag acg gcc	720
Thr Asn Phe Ala Gly Leu Lys Lys Met Ala Ala Ala Cys Gln Thr Ala	
225 230 235 240	
atc ccg tcc tgg ctg ggg aac ctg ttc gac ggg ctg gag aac gac gcg	768
Ile Pro Ser Trp Leu Gly Asn Leu Phe Asp Gly Leu Glu Asn Asp Ala	
245 250 255	
gag acc cgc cgc ctg atc gcc tgt tgc gtg gcc gcc gag atg tgc gcc	816
Glu Thr Arg Arg Leu Ile Ala Cys Ser Val Ala Ala Glu Met Cys Ala	
260 265 270	
aag ctg cag gaa cag ggt ttc gag gac ttc cac ttc tac acc ctg aac	864
Lys Leu Gln Glu Gln Gly Phe Glu Asp Phe His Phe Tyr Thr Leu Asn	
275 280 285	
cgg gcc gat ctc gtt tac gcc atc tgc cgt gtg ctg ggc gtg cgc gag	912
Arg Ala Asp Leu Val Tyr Ala Ile Cys Arg Val Leu Gly Val Arg Glu	
290 295 300	
atc tgc ccc gcc gct tgc gag gtc gcc gca tga	945
Ile Ser Pro Ala Ala Ser Glu Val Ala Ala	
305 310	

<210> 40

<211> 314

<212> PRT

<213> *Caulobacter crescentus*

<400> 40

Met Thr Leu Pro Pro Thr Arg Arg Val Ile Gly Pro Val Ala Arg Ala
 1 5 10 15
 Gly Glu Arg Thr Gly Arg Pro Arg Val Ser Phe Glu Phe Phe Pro Pro
 20 25 30
 Lys Thr Pro Gln Met Glu Glu Ser Leu Trp Gln Ala Ile Thr Arg Leu
 35 40 45
 Ala Pro Leu Asp Pro Ala Phe Val Ser Val Thr Tyr Gly Ala Gly Gly
 50 55 60
 Ser Thr Arg Glu Arg Thr His Arg Thr Val Lys Arg Ile Leu Asp Glu
 65 70 75 80
 Thr Ser Leu Lys Pro Ala Ala His Leu Thr Cys Val Gly Ala Ser Arg
 85 90 95
 Glu Glu Val Asp Glu Val Ile Arg Glu Tyr Trp Glu Thr Gly Val Arg
 100 105 110
 His Ile Val Ser Leu Arg Gly Asp Pro Pro Pro Gly Glu Gly Gly Ile
 115 120 125
 Gly Gly Val Tyr Val Pro Arg Ala Asp Gly Tyr Ala Asn Ala Thr Glu
 130 135 140
 Leu Thr Lys Ala Val Arg Ala Ile Ala Pro Phe Glu Val Leu Val Gly
 145 150 155 160
 Val Tyr Pro Glu Lys His Pro Glu Ser Pro Ser Leu Glu His Asp Ile
 165 170 175
 Asp Val Leu Lys Gln Lys Val Asp Ala Gly Ala Thr Leu Gly Ile Ser
 180 185 190
 Gln Phe Phe Phe Asp Leu Asp Ala Phe Leu Arg Phe Val Asp Lys Val
 195 200 205
 Arg Ala Ala Gly Ile Thr Ile Pro Ile Val Pro Gly Ile Met Pro Val
 210 215 220
 Thr Asn Phe Ala Gly Leu Lys Lys Met Ala Ala Ala Cys Gln Thr Ala
 225 230 235 240
 Ile Pro Ser Trp Leu Gly Asn Leu Phe Asp Gly Leu Glu Asn Asp Ala
 245 250 255
 Glu Thr Arg Arg Leu Ile Ala Cys Ser Val Ala Ala Glu Met Cys Ala
 260 265 270
 Lys Leu Gln Glu Gln Gly Phe Glu Asp Phe His Phe Tyr Thr Leu Asn
 275 280 285
 Arg Ala Asp Leu Val Tyr Ala Ile Cys Arg Val Leu Gly Val Arg Glu
 290 295 300
 Ile Ser Pro Ala Ala Ser Glu Val Ala Ala
 305 310

<210> 41
 <211> 885
 <212> DNA
 <213> Actinobacillus actinomycetemcomitans

<220>
 <221> CDS
 <222> (1)..(882)

<223> RAB00260

<400> 41

atg agt tac gca aaa gaa att gat aat cta aat caa cat tta gct gat	48
Met Ser Tyr Ala Lys Glu Ile Asp Asn Leu Asn Gln His Leu Ala Asp	
1 5 10 15	
tta aac ggc aaa att aat gtc tct ttt gaa ttt ttc ccg ccg aaa agt	96
Leu Asn Gly Lys Ile Asn Val Ser Phe Glu Phe Phe Pro Pro Lys Ser	
20 25 30	
aaa aaa atg gaa aat ctt ctg tgg gaa tcc atc cat cgc tta aaa gtg	144
Glu Lys Met Glu Asn Leu Leu Trp Glu Ser Ile His Arg Leu Lys Val	
35 40 45	
cta aaa ccg aaa ttt gta tcc gtg act tac ggc gcc aat tcc ggc gag	192
Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser Gly Glu	
50 55 60	
cgt gaa cgc act cac ggg gtg gtg aaa cgc att aag cag gaa acc ggt	240
Arg Glu Arg Thr His Gly Val Val Lys Arg Ile Lys Gln Glu Thr Gly	
65 70 75 80	
ctg gaa gct gcg ccg cat tta acc ggt att gac gct acc tcg gac gaa	288
Leu Glu Ala Ala Pro His Leu Thr Gly Ile Asp Ala Thr Ser Asp Glu	
85 90 95	
ttg cgt cgc att gcc aaa ggt tat tgg gat agc ggc att cgt cgc att	336
Leu Arg Arg Ile Ala Lys Gly Tyr Trp Asp Ser Gly Ile Arg Arg Ile	
100 105 110	
ctg gca ctg cgc ggt gac gag ccg aaa ggc tac gag aaa aaa cca ttt	384
Val Ala Leu Arg Gly Asp Glu Pro Lys Gly Tyr Glu Lys Lys Pro Phe	
115 120 125	
tat gcc gcc gat tta gta gca tta tta cgt gac gta tca gat ttt gat	432
Tyr Ala Ala Asp Leu Val Ala Leu Leu Arg Asp Val Ser Asp Phe Asp	
130 135 140	
att tcc gtg gcg gca tac cct gag gtt cat ccg gaa gcc aaa tcg gcg	480
Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys Ser Ala	
145 150 155 160	
caa gcg gat tta att aat tta aaa cgt aaa att gat gcc ggt gcc aat	528
Gln Ala Asp Leu Ile Asn Leu Lys Arg Lys Ile Asp Ala Gly Ala Asn	
165 170 175	
cat gtg atc aca caa ttc ttt ttc gat att gac agc tat ctg cgg ttc	576
His Val Ile Thr Gln Phe Phe Phe Asp Ile Asp Ser Tyr Leu Arg Phe	
180 185 190	
cgc gat cgc tgc gcg tct atc ggt att gat gca gaa atc gtg ccg ggg	624

Arg Asp Arg Cys Ala Ser Ile Gly Ile Asp Ala Glu Ile Val Pro Gly
 195 200 205

att ctg ccg gtg acc aac ttc aaa caa tta caa aaa atg gca gca atc 672
 Ile Leu Pro Val Thr Asn Phe Lys Gln Leu Gln Lys Met Ala Ala Ile
 210 215 220

act aat gtg aaa att cca gct tgg atg agc aaa atg tat gaa ggc ttg 720
 Thr Asn Val Lys Ile Pro Ala Trp Met Ser Lys Met Tyr Glu Gly Leu
 225 230 235 240

gat gat gac caa acc acc cgc aat ctg gtg gcg gcg agc atc gcc atg 768
 Asp Asp Asp Gln Thr Thr Arg Asn Leu Val Ala Ala Ser Ile Ala Met
 245 250 255

gac atg gtg cgt gta ctg tcc cgc gaa ggg gta aaa gac ttt cat ttc 816
 Asp Met Val Arg Val Leu Ser Arg Glu Gly Val Lys Asp Phe His Phe
 260 265 270

tac acc ctg aat cgc agt gaa ctc acc tat gct att tgc cac acg tta 864
 Tyr Thr Leu Asn Arg Ser Glu Leu Thr Tyr Ala Ile Cys His Thr Leu
 275 280 285

ggc att cgt ccg agt ttg taa 885
 Gly Ile Arg Pro Ser Leu
 290

<210> 42

<211> 294

<212> PRT

<213> Actinobacillus actinomycetemcomitans

<400> 42

Met Ser Tyr Ala Lys Glu Ile Asp Asn Leu Asn Gln His Leu Ala Asp
 1 5 10 15

Leu Asn Gly Lys Ile Asn Val Ser Phe Glu Phe Phe Pro Pro Lys Ser
 20 25 30

Glu Lys Met Glu Asn Leu Leu Trp Glu Ser Ile His Arg Leu Lys Val
 35 40 45

Leu Lys Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser Gly Glu
 50 55 60

Arg Glu Arg Thr His Gly Val Val Lys Arg Ile Lys Gln Glu Thr Gly
 65 70 75 80

Leu Glu Ala Ala Pro His Leu Thr Gly Ile Asp Ala Thr Ser Asp Glu
 85 90 95

Leu Arg Arg Ile Ala Lys Gly Tyr Trp Asp Ser Gly Ile Arg Arg Ile
 100 105 110

Val Ala Leu Arg Gly Asp Glu Pro Lys Gly Tyr Glu Lys Lys Pro Phe
 115 120 125

Tyr Ala Ala Asp Leu Val Ala Leu Leu Arg Asp Val Ser Asp Phe Asp
 130 135 140

Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys Ser Ala
 145 150 155 160

Gln Ala Asp Leu Ile Asn Leu Lys Arg Lys Ile Asp Ala Gly Ala Asn
 165 170 175
 His Val Ile Thr Gln Phe Phe Phe Asp Ile Asp Ser Tyr Leu Arg Phe
 180 185 190
 Arg Asp Arg Cys Ala Ser Ile Gly Ile Asp Ala Glu Ile Val Pro Gly
 195 200 205
 Ile Leu Pro Val Thr Asn Phe Lys Gln Leu Gln Lys Met Ala Ala Ile
 210 215 220
 Thr Asn Val Lys Ile Pro Ala Trp Met Ser Lys Met Tyr Glu Gly Leu
 225 230 235 240
 Asp Asp Asp Gln Thr Thr Arg Asn Leu Val Ala Ala Ser Ile Ala Met
 245 250 255
 Asp Met Val Arg Val Leu Ser Arg Glu Gly Val Lys Asp Phe His Phe
 260 265 270
 Thr Thr Leu Asn Arg Ser Glu Leu Thr Tyr Ala Ile Cys His Thr Leu
 275 280 285
 Gly Ile Arg Pro Ser Leu
 290

<210> 43
 <211> 867
 <212> DNA
 <213> Rhodobacter

<220>
 <221> CDS
 <222> (1) .. (864)
 <223> RRC03981

<400> 43
 atg acc acg ccg cat gtc agc ttt gaa ttc ttc ccg ccg cag acg ctc 48
 Met Thr Thr Pro His Val Ser Phe Glu Phe Phe Pro Pro Gln Thr Leu
 1 5 10 15
 gac gcc tcg ttc cgg ctg tgg gag acg gcg cag gtt ctg gcg ccg ctc 96
 Asp Ala Ser Phe Arg Leu Trp Glu Thr Ala Gln Val Leu Ala Pro Leu
 20 25 30
 aag ccc ggc ttc gtc tcg gtc acc tat ggc gcg ggc ggc acc acc cgc 144
 Lys Pro Gly Phe Val Ser Val Thr Tyr Gly Ala Gly Gly Thr Thr Arg
 35 40 45
 aag ctg acg cat gag gcc gtg gcg gcg atc cac aag aat tac ggc ctg 192
 Lys Leu Thr His Glu Ala Val Ala Ala Ile His Lys Asn Tyr Gly Leu
 50 55 60
 aac gtc gcc gcg cat ctg acc tgc gtc gat gcg acc ccg gcc gaa acg 240
 Asn Val Ala Ala His Leu Thr Cys Val Asp Ala Thr Arg Ala Glu Thr
 65 70 75 80
 caa gag atc atc gac gcc tat gcc gag gct ggc gtc acc gag att gtc 288
 Gln Glu Ile Ile Asp Ala Tyr Ala Glu Ala Gly Val Thr Glu Ile Val
 85 90 95

gcg ctg cgc ggt gat ccg ccg aaa ggc gcc gcc cgc ttc acg ccg cat	336
Ala Leu Arg Gly Asp Pro Pro Lys Gly Ala Ala Arg Phe Thr Pro His	
100 105 110	
ccg gac ggg ttt gcc tcc tgc gtg gac ctc atc gaa tgg ctg gcg cgg	384
Pro Asp Gly Phe Ala Ser Ser Val Asp Leu Ile Glu Trp Leu Ala Arg	
115 120 125	
gac ggc cgc ttc acg ctg cgc tgc ggc gcc tat ccg gaa ccg cat ccg	432
Asp Gly Arg Phe Thr Leu Arg Cys Gly Ala Tyr Pro Glu Pro His Pro	
130 135 140	
gaa gcc gcc gac acg ctg gcc gac gtg cgc tgg ctg aaa cgc aaa tgc	480
Glu Ala Ala Asp Thr Leu Ala Asp Val Arg Trp Leu Lys Arg Lys Cys	
145 150 155 160	
gag gcg ggg gcg acc tgc gcg atc acg caa ttc ttc ttt gaa gcc gag	528
Glu Ala Gly Ala Thr Ser Ala Ile Thr Gln Phe Phe Phe Glu Ala Glu	
165 170 175	
c ttc ttc cgc ttc cgc gac gcc tgc gtg aag gaa ggg atc acc gcc	576
Ar Phe Phe Arg Phe Arg Asp Ala Cys Val Lys Glu Gly Ile Thr Ala	
180 185 190	
aag atc atc ccg ggc atc ctg ccg atc cag tcc tgg aaa ggc gcc aag	624
Lys Ile Ile Pro Gly Ile Leu Pro Ile Gln Ser Trp Lys Gly Ala Lys	
195 200 205	
agc ttt gcg cag cgc tgc ggc acc tgc atc ccg acc tgg gtc gaa gag	672
Ser Phe Ala Gln Arg Cys Gly Thr Ser Ile Pro Thr Trp Val Glu Glu	
210 215 220	
gcc ttt gac cat gcg atc cgc gac gac cgc gaa cag ctg ctg gcc acg	720
Ala Phe Asp His Ala Ile Arg Asp Asp Arg Glu Gln Leu Leu Ala Thr	
225 230 235 240	
gcg ctg tgc acg gag ctc tgc gac aac ctg atc gcg ggc ggg gtg gag	768
Ala Leu Cys Thr Glu Leu Cys Asp Asn Leu Ile Ala Gly Gly Val Glu	
245 250 255	
at ctg cat ttc tac acg ctg aac cgg ccg cag atg acc cgc gat gtc	816
asp Leu His Phe Tyr Thr Leu Asn Arg Pro Gln Met Thr Arg Asp Val	
260 265 270	
tgc cat gcg ctg ggc gtc aac ccg ggt gtg gtg ctg gaa aac gtc gcc	864
Cys His Ala Leu Gly Val Asn Pro Gly Val Val Leu Glu Asn Val Ala	
275 280 285	
tga	867

<210> 44

<211> 288

<212> PRT

<213> Rhodobacter

<400> 44

Met Thr Thr Pro His Val Ser Phe Glu Phe Phe Pro Pro Gln Thr Leu
1 5 10 15

Asp Ala Ser Phe Arg Leu Trp Glu Thr Ala Gln Val Leu Ala Pro Leu
20 25 30

```

Lys Pro Gly Phe Val Ser Val Thr Tyr Gly Ala Gly Gly Thr Thr Arg
    35                                40                                45

Lys Leu Thr His Glu Ala Val Ala Ala Ile His Lys Asn Tyr Gly Leu
    50                                55                                60

Asn Val Ala Ala His Leu Thr Cys Val Asp Ala Thr Arg Ala Glu Thr
    65                                70                                75                                80

Gln Glu Ile Ile Asp Ala Tyr Ala Glu Ala Gly Val Thr Glu Ile Val
    85                                90                                95

Ala Leu Arg Gly Asp Pro Pro Lys Gly Ala Ala Arg Phe Thr Pro His
    100                               105                               110

Pro Asp Gly Phe Ala Ser Ser Val Asp Leu Ile Glu Trp Leu Ala Arg
    115                               120                               125

Asp Gly Arg Phe Thr Leu Arg Cys Gly Ala Tyr Pro Glu Pro His Pro
    130                               135                               140

Leu Ala Ala Asp Thr Leu Ala Asp Val Arg Trp Leu Lys Arg Lys Cys
    145                               150                               155                               160

Glu Ala Gly Ala Thr Ser Ala Ile Thr Gln Phe Phe Phe Glu Ala Glu
    165                               170                               175

Thr Phe Phe Arg Phe Arg Asp Ala Cys Val Lys Glu Gly Ile Thr Ala
    180                               185                               190

Lys Ile Ile Pro Gly Ile Leu Pro Ile Gln Ser Trp Lys Gly Ala Lys
    195                               200                               205

Ser Phe Ala Gln Arg Cys Gly Thr Ser Ile Pro Thr Trp Val Glu Glu
    210                               215                               220

Ala Phe Asp His Ala Ile Arg Asp Asp Arg Glu Gln Leu Leu Ala Thr
    225                               230                               235                               240

Ala Leu Cys Thr Glu Leu Cys Asp Asn Leu Ile Ala Gly Gly Val Glu
    245                               250                               255

Asp Leu His Phe Tyr Thr Leu Asn Arg Pro Gln Met Thr Arg Asp Val
    260                               265                               270

Cys His Ala Leu Gly Val Asn Pro Gly Val Val Leu Glu Asn Val Ala
    275                               280                               285

```

<210> 45

<211> 879

<212> DNA

<213> Neisseria meningitidis ser. A

<220>

<221> CDS

<222> (1) .. (876)

<223> RNM00812

<400> 45

atg aat tac gca aaa gaa atc aat gcg tta aat aac agc ctt tcc gat	48
Met Asn Tyr Ala Lys Glu Ile Asn Ala Leu Asn Asn Ser Leu Ser Asp	
1 5 10 15	
ttg aaa ggc gac atc aac gtt tcg ttt gaa ttt ttt cca ccg aaa aac	96
Leu Lys Gly Asp Ile Asn Val Ser Phe Glu Phe Phe Pro Pro Lys Asn	
20 25 30	
gag caa atg gaa acg atg ctg tgg gat tcc atc cac cgt ctg caa acc	144
Glu Gln Met Glu Thr Met Leu Trp Asp Ser Ile His Arg Leu Gln Thr	
35 40 45	
ctg cat ccc aag ttc gta tcc gta acc tac ggc gca aac tcc ggc gaa	192
Leu His Pro Lys Phe Val Ser Val Thr Tyr Gly Ala Asn Ser Gly Glu	
50 55 60	
cgc gac cgc acg cac ggc atc gtc aaa cgc atc aaa cag gaa acc ggc	240
Arg Asp Arg Thr His Gly Ile Val Lys Arg Ile Lys Gln Glu Thr Gly	
65 70 75 80	
g gaa gca gca ccg cac ctg acc ggc atc gac gca tcc ccc gac gaa	288
Leu Glu Ala Ala Pro His Leu Thr Gly Ile Asp Ala Ser Pro Asp Glu	
85 90 95	
ttg cgc caa atc gcc aaa gac tat tgg gac agc ggc atc cgc cgc att	336
Leu Arg Gln Ile Ala Lys Asp Tyr Trp Asp Ser Gly Ile Arg Arg Ile	
100 105 110	
gtc gcc ctg cgt ggc gac gag ccg ccc ggt tat gag aaa aaa ccg ttt	384
Val Ala Leu Arg Gly Asp Glu Pro Pro Gly Tyr Glu Lys Lys Pro Phe	
115 120 125	
tac gcc gaa gac ttg gtt aag cta tta cgc tcc gtc gcc gac ttc gac	432
Tyr Ala Glu Asp Leu Val Lys Leu Leu Arg Ser Val Ala Asp Phe Asp	
130 135 140	
atc tct gtg gcg gca tat ccc gaa gtg cat ccc gaa gcc aaa tcc gca	480
Ile Ser Val Ala Ala Tyr Pro Glu Val His Pro Glu Ala Lys Ser Ala	
145 150 155 160	
aa gcc gat ctg att aat ctg aag cgc aaa atc gat gcg ggt gca aac	528
Gln Ala Asp Leu Ile Asn Leu Lys Arg Lys Ile Asp Ala Gly Ala Asn	
165 170 175	
cac gtc atc acc caa ttt ttc ttt gac gta gaa cgc tac ctg cgc ttc	576
His Val Ile Thr Gln Phe Phe Phe Asp Val Glu Arg Tyr Leu Arg Phe	
180 185 190	
cgc gac cgc tgc gtg atg ttg ggt atc gat gtg gaa atc gtc cct ggt	624
Arg Asp Arg Cys Val Met Leu Gly Ile Asp Val Glu Ile Val Pro Gly	
195 200 205	
att ttg cct gtt acc aac ttc aag cag ctc ggc aaa atg gcg caa gta	672
Ile Leu Pro Val Thr Asn Phe Lys Gln Leu Gly Lys Met Ala Gln Val	
210 215 220	
acc aac gtc aaa atc cca agc tgg ctg tcg caa atg tat gaa ggt ttg	720
Thr Asn Val Lys Ile Pro Ser Trp Leu Ser Gln Met Tyr Glu Gly Leu	
225 230 235 240	
gac gac gac caa ggc acg cgc aac ctc gtc gcc gcc agt atc gcc atc	768
Asp Asp Asp Gln Gly Thr Arg Asn Leu Val Ala Ala Ser Ile Ala Ile	

245										250					255					
gat	atg	gtc	aaa	gtc	ctg	tcc	cgc	gaa	ggc	gtg	aaa	gat	ttc	cac	ttc	816				
Asp	Met	Val	Lys	Val	Leu	Ser	Arg	Glu	Gly	Val	Lys	Asp	Phe	His	Phe					
			260						265				270							
tac	acg	ctc	aac	cgc	agc	gag	ctg	act	tac	gcc	atc	tgc	cat	att	tta	864				
Tyr	Thr	Leu	Asn	Arg	Ser	Glu	Leu	Thr	Tyr	Ala	Ile	Cys	His	Ile	Leu					
		275					280					285								
ggc	gtg	cgc	cct	taa												879				
Gly	Val	Arg	Pro																	
			290																	

<210> 46

<211> 292

<212> PRT

<213> Neisseria meningitidis ser. A

<400> 46

t	Asn	Tyr	Ala	Lys	Glu	Ile	Asn	Ala	Leu	Asn	Asn	Ser	Leu	Ser	Asp	
1				5					10					15		
Leu	Lys	Gly	Asp	Ile	Asn	Val	Ser	Phe	Glu	Phe	Phe	Pro	Pro	Lys	Asn	
			20					25					30			
Glu	Gln	Met	Glu	Thr	Met	Leu	Trp	Asp	Ser	Ile	His	Arg	Leu	Gln	Thr	
		35					40					45				
Leu	His	Pro	Lys	Phe	Val	Ser	Val	Thr	Tyr	Gly	Ala	Asn	Ser	Gly	Glu	
	50					55					60					
Arg	Asp	Arg	Thr	His	Gly	Ile	Val	Lys	Arg	Ile	Lys	Gln	Glu	Thr	Gly	
	65				70					75					80	
Leu	Glu	Ala	Ala	Pro	His	Leu	Thr	Gly	Ile	Asp	Ala	Ser	Pro	Asp	Glu	
				85					90					95		
Leu	Arg	Gln	Ile	Ala	Lys	Asp	Tyr	Trp	Asp	Ser	Gly	Ile	Arg	Arg	Ile	
			100					105					110			
al	Ala	Leu	Arg	Gly	Asp	Glu	Pro	Pro	Gly	Tyr	Glu	Lys	Lys	Pro	Phe	
		115					120					125				
Tyr	Ala	Glu	Asp	Leu	Val	Lys	Leu	Leu	Arg	Ser	Val	Ala	Asp	Phe	Asp	
		130				135					140					
Ile	Ser	Val	Ala	Ala	Tyr	Pro	Glu	Val	His	Pro	Glu	Ala	Lys	Ser	Ala	
	145				150					155					160	
Gln	Ala	Asp	Leu	Ile	Asn	Leu	Lys	Arg	Lys	Ile	Asp	Ala	Gly	Ala	Asn	
			165					170					175			
His	Val	Ile	Thr	Gln	Phe	Phe	Phe	Asp	Val	Glu	Arg	Tyr	Leu	Arg	Phe	
		180						185					190			
Arg	Asp	Arg	Cys	Val	Met	Leu	Gly	Ile	Asp	Val	Glu	Ile	Val	Pro	Gly	
		195					200					205				
Ile	Leu	Pro	Val	Thr	Asn	Phe	Lys	Gln	Leu	Gly	Lys	Met	Ala	Gln	Val	
	210					215					220					

Thr Asn Val Lys Ile Pro Ser Trp Leu Ser Gln Met Tyr Glu Gly Leu
 225 230 235 240
 Asp Asp Asp Gln Gly Thr Arg Asn Leu Val Ala Ala Ser Ile Ala Ile
 245 250 255
 Asp Met Val Lys Val Leu Ser Arg Glu Gly Val Lys Asp Phe His Phe
 260 265 270
 Tyr Thr Leu Asn Arg Ser Glu Leu Thr Tyr Ala Ile Cys His Ile Leu
 275 280 285
 Gly Val Arg Pro
 290

<210> 47
 <211> 849
 <212> DNA
 <213> Campylobacter jejuni

20>
 21> CDS
 <222> (1) .. (846)
 <223> RCJ02911

<400> 47
 atg tgt agt ttt tct ttt gaa gtt ttt cca cca aga aag gat gaa aat 48
 Met Cys Ser Phe Ser Phe Glu Val Phe Pro Pro Arg Lys Asp Glu Asn
 1 5 10 15
 atc aaa aat ctt cat gct atc tta gat gat tta ggg caa tta agc cct 96
 Ile Lys Asn Leu His Ala Ile Leu Asp Asp Leu Gly Gln Leu Ser Pro
 20 25 30
 aat ttt atc agc gta acc ttt gga gct gga ggc tct att aac tca caa 144
 Asn Phe Ile Ser Val Thr Phe Gly Ala Gly Gly Ser Ile Asn Ser Gln
 35 40 45
 aat act tta gaa gtt gca agc tta atc cag gaa gaa tat caa att cct 192
 Asn Thr Leu Glu Val Ala Ser Leu Ile Gln Glu Glu Tyr Gln Ile Pro
 50 55 60
 agc ata gta cat tta cct tgc atc cat tct agt aaa gaa aaa atc act 240
 Ser Ile Val His Leu Pro Cys Ile His Ser Ser Lys Glu Lys Ile Thr
 65 70 75 80
 cag ata ctt caa aaa tgc aaa gaa aaa aat ctt aat caa att ctt gcc 288
 Gln Ile Leu Gln Lys Cys Lys Glu Lys Asn Leu Asn Gln Ile Leu Ala
 85 90 95
 cta aga ggc gat ata tgt gaa aat tta aaa aaa agc aaa gat ttt tct 336
 Leu Arg Gly Asp Ile Cys Glu Asn Leu Lys Lys Ser Lys Asp Phe Ser
 100 105 110
 tat gct agt gat tta att tct ttt ata aaa aaa caa gaa tac ttt gaa 384
 Tyr Ala Ser Asp Leu Ile Ser Phe Ile Lys Lys Gln Glu Tyr Phe Glu
 115 120 125
 att tat gcc gca tgc tat ccc gaa aaa cat aat gaa tct aaa aat ttc 432
 Ile Tyr Ala Ala Cys Tyr Pro Glu Lys His Asn Glu Ser Lys Asn Phe
 130 135 140

atc gag gat ata cac cat ctt aaa act aag gta aat gca gga aca gat 480
 Ile Glu Asp Ile His His Leu Lys Thr Lys Val Asn Ala Gly Thr Asp
 145 150 155 160
 aag ctc att act caa ctt ttt tac gat aat gaa gat ttt tat act ttt 528
 Lys Leu Ile Thr Gln Leu Phe Tyr Asp Asn Glu Asp Phe Tyr Thr Phe
 165 170 175
 aaa caa aat tgt gct tta gca gat att gac ata cct att tac gca ggt 576
 Lys Gln Asn Cys Ala Leu Ala Asp Ile Asp Ile Pro Ile Tyr Ala Gly
 180 185 190
 att atg cct att act aac aaa aga cag gtt tta aaa att tct caa ctt 624
 Ile Met Pro Ile Thr Asn Lys Arg Gln Val Leu Lys Ile Ser Gln Leu
 195 200 205
 tgc gga gct aaa atc cct cct aaa ttt gtt aaa att tta gaa aaa tat 672
 Cys Gly Ala Lys Ile Pro Pro Lys Phe Val Lys Ile Leu Glu Lys Tyr
 210 215 220
 aat aat act ttg gct tta gaa gat gca ggt atc gcg tat gct tgc 720
 Asn Asn Thr Leu Ala Leu Glu Asp Ala Gly Ile Ala Tyr Ala Cys
 225 230 235 240
 gat caa att gtc gat tta atc aca agt ggt gta gat gga att cat ctt 768
 Asp Gln Ile Val Asp Leu Ile Thr Ser Gly Val Asp Gly Ile His Leu
 245 250 255
 tat act atg aat aaa tcc aaa gcg gct att aaa att tat gaa gct gta 816
 Tyr Thr Met Asn Lys Ser Lys Ala Ala Ile Lys Ile Tyr Glu Ala Val
 260 265 270
 aag cat ttg ctt aaa gaa gag ctt cat gct tag 849
 Lys His Leu Leu Lys Glu Glu Leu His Ala
 275 280

<210> 48

<211> 282

<212> PRT

<213> Campylobacter jejuni

<400> 48

Met Cys Ser Phe Ser Phe Glu Val Phe Pro Pro Arg Lys Asp Glu Asn
 1 5 10 15
 Ile Lys Asn Leu His Ala Ile Leu Asp Asp Leu Gly Gln Leu Ser Pro
 20 25 30
 Asn Phe Ile Ser Val Thr Phe Gly Ala Gly Gly Ser Ile Asn Ser Gln
 35 40 45
 Asn Thr Leu Glu Val Ala Ser Leu Ile Gln Glu Glu Tyr Gln Ile Pro
 50 55 60
 Ser Ile Val His Leu Pro Cys Ile His Ser Ser Lys Glu Lys Ile Thr
 65 70 75 80
 Gln Ile Leu Gln Lys Cys Lys Glu Lys Asn Leu Asn Gln Ile Leu Ala
 85 90 95
 Leu Arg Gly Asp Ile Cys Glu Asn Leu Lys Lys Ser Lys Asp Phe Ser

100	105	110
Tyr Ala Ser Asp Leu Ile Ser Phe Ile Lys Lys Gln Glu Tyr Phe Glu 115 120 125		
Ile Tyr Ala Ala Cys Tyr Pro Glu Lys His Asn Glu Ser Lys Asn Phe 130 135 140		
Ile Glu Asp Ile His His Leu Lys Thr Lys Val Asn Ala Gly Thr Asp 145 150 155 160		
Lys Leu Ile Thr Gln Leu Phe Tyr Asp Asn Glu Asp Phe Tyr Thr Phe 165 170 175		
Lys Gln Asn Cys Ala Leu Ala Asp Ile Asp Ile Pro Ile Tyr Ala Gly 180 185 190		
Ile Met Pro Ile Thr Asn Lys Arg Gln Val Leu Lys Ile Ser Gln Leu 195 200 205		
Cys Gly Ala Lys Ile Pro Pro Lys Phe Val Lys Ile Leu Glu Lys Tyr 210 215 220		
Asn Asn Thr Leu Ala Leu Glu Asp Ala Gly Ile Ala Tyr Ala Cys 225 230 235 240		
Asp Gln Ile Val Asp Leu Ile Thr Ser Gly Val Asp Gly Ile His Leu 245 250 255		
Tyr Thr Met Asn Lys Ser Lys Ala Ala Ile Lys Ile Tyr Glu Ala Val 260 265 270		
Lys His Leu Leu Lys Glu Glu Leu His Ala 275 280		

<210> 49

<211> 852

<212> DNA

<213> Lactococcus lactis

<220>

<221> CDS

<222> (1)..(849)

<223> AAK05352

<400> 49

atg aca agt aat tcc aaa att ctt tct ttt gaa gtt ttt cca cct aca	48
Met Thr Ser Asn Ser Lys Ile Leu Ser Phe Glu Val Phe Pro Pro Thr	
1 5 10 15	

act caa att gga agt acc aac ttg gta aag acc ttg gat agc cta aga	96
Thr Gln Ile Gly Ser Thr Asn Leu Val Lys Thr Leu Asp Ser Leu Arg	
20 25 30	

act ctc tcg cca gat ttt atc agt gta act tgt agt aac aat aat tat	144
Thr Leu Ser Pro Asp Phe Ile Ser Val Thr Cys Ser Asn Asn Asn Tyr	
35 40 45	

gat aat att gga gat aca act ata aag ttt gct gat tat gta aac aat	192
Asp Asn Ile Gly Asp Thr Thr Ile Lys Phe Ala Asp Tyr Val Asn Asn	
50 55 60	

aca cta gat att cca gcg gtt gct cat tta cct gcc gct tat tta gat	240
Thr Leu Asp Ile Pro Ala Val Ala His Leu Pro Ala Ala Tyr Leu Asp	
65 70 75 80	
aaa gct caa gtg atc gaa att ttg gaa cgg tta aaa gat aaa caa atc	288
Lys Ala Gln Val Ile Glu Ile Leu Glu Arg Leu Lys Asp Lys Gln Ile	
85 90 95	
aaa aaa att ctt gct tta aga ggt gat atc agc gat gaa ccg atg aaa	336
Lys Lys Ile Leu Ala Leu Arg Gly Asp Ile Ser Asp Glu Pro Met Lys	
100 105 110	
gat gat ttt aaa ttt gca agt gat ttg gtt aaa ttt atc aaa gat tat	384
Asp Asp Phe Lys Phe Ala Ser Asp Leu Val Lys Phe Ile Lys Asp Tyr	
115 120 125	
gat gat agt ttt gaa gtt tta ggt gct tgc tac ccc gat att cat ccc	432
Asp Asp Ser Phe Glu Val Leu Gly Ala Cys Tyr Pro Asp Ile His Pro	
130 135 140	
gaa tca gta aat cga gtg agt gat ttt cat tat ctg aaa gaa aaa gta	480
Val Ser Val Asn Arg Val Ser Asp Phe His Tyr Leu Lys Glu Lys Val	
150 155 160	
gat gct ggt tgt gac aga tta atc acg caa cta ttt ttt gat aat gat	528
Asp Ala Gly Cys Asp Arg Leu Ile Thr Gln Leu Phe Phe Asp Asn Asp	
165 170 175	
agt ttc tat gat ttt caa gaa cga tgc gca att gct gag ata aat act	576
Ser Phe Tyr Asp Phe Gln Glu Arg Cys Ala Ile Ala Glu Ile Asn Thr	
180 185 190	
ccg ata ttc gcc gga ata atg cca gta atc aat cga aat caa att ctt	624
Pro Ile Phe Ala Gly Ile Met Pro Val Ile Asn Arg Asn Gln Ile Leu	
195 200 205	
cgt cta tta aaa aat tgt aat acg cca tta cca gca aaa ttc att aga	672
Arg Leu Leu Lys Asn Cys Asn Thr Pro Leu Pro Ala Lys Phe Ile Arg	
210 215 220	
ata ctc gaa aaa tat gaa cat aat ctt atc gct tta agg gat gct gga	720
Ile Leu Glu Lys Tyr Glu His Asn Leu Ile Ala Leu Arg Asp Ala Gly	
230 235 240	
att gct tac gcc atc gat caa atc gtt gat tta gta aca gag gat gtt	768
Ile Ala Tyr Ala Ile Asp Gln Ile Val Asp Leu Val Thr Glu Asp Val	
245 250 255	
gct gga att cac ctc tat acg atg aat aat gca aat acg gca cac tcc	816
Ala Gly Ile His Leu Tyr Thr Met Asn Asn Ala Asn Thr Ala His Ser	
260 265 270	
atc cat gct tca att tct tct tta ttt acc ttt tga	852
Ile His Ala Ser Ile Ser Ser Leu Phe Thr Phe	
275 280	

<210> 50

<211> 283

<212> PRT

<213> Lactococcus lactis.

<400> 50

Met Thr Ser Asn Ser Lys Ile Leu Ser Phe Glu Val Phe Pro Pro Thr
 1 5 10 15

Thr Gln Ile Gly Ser Thr Asn Leu Val Lys Thr Leu Asp Ser Leu Arg
 20 25 30

Thr Leu Ser Pro Asp Phe Ile Ser Val Thr Cys Ser Asn Asn Asn Tyr
 35 40 45

Asp Asn Ile Gly Asp Thr Thr Ile Lys Phe Ala Asp Tyr Val Asn Asn
 50 55 60

Thr Leu Asp Ile Pro Ala Val Ala His Leu Pro Ala Ala Tyr Leu Asp
 65 70 75 80

Lys Ala Gln Val Ile Glu Ile Leu Glu Arg Leu Lys Asp Lys Gln Ile
 85 90 95

Lys Lys Ile Leu Ala Leu Arg Gly Asp Ile Ser Asp Glu Pro Met Lys
 100 105 110

Asp Phe Lys Phe Ala Ser Asp Leu Val Lys Phe Ile Lys Asp Tyr
 115 120 125

Asp Asp Ser Phe Glu Val Leu Gly Ala Cys Tyr Pro Asp Ile His Pro
 130 135 140

Glu Ser Val Asn Arg Val Ser Asp Phe His Tyr Leu Lys Glu Lys Val
 145 150 155 160

Asp Ala Gly Cys Asp Arg Leu Ile Thr Gln Leu Phe Phe Asp Asn Asp
 165 170 175

Ser Phe Tyr Asp Phe Gln Glu Arg Cys Ala Ile Ala Glu Ile Asn Thr
 180 185 190

Pro Ile Phe Ala Gly Ile Met Pro Val Ile Asn Arg Asn Gln Ile Leu
 195 200 205

Arg Leu Leu Lys Asn Cys Asn Thr Pro Leu Pro Ala Lys Phe Ile Arg
 210 215 220

Leu Glu Lys Tyr Glu His Asn Leu Ile Ala Leu Arg Asp Ala Gly
 225 230 235 240

Ile Ala Tyr Ala Ile Asp Gln Ile Val Asp Leu Val Thr Glu Asp Val
 245 250 255

Ala Gly Ile His Leu Tyr Thr Met Asn Asn Ala Asn Thr Ala His Ser
 260 265 270

Ile His Ala Ser Ile Ser Ser Leu Phe Thr Phe
 275 280

<210> 51

<211> 891

<212> DNA

<213> *Prochlorococcus maritima*

<220>

<221> CDS

M/43126

MetF

<222> (1)..(888)

<223> RCK01602

<400> 51

ttg	aaa	tca	aaa	ctt	cag	caa	act	tta	gaa	aag	aat	tca	aaa	gta	att	48
Leu	Lys	Ser	Lys	Leu	Gln	Gln	Thr	Leu	Glu	Lys	Asn	Ser	Lys	Val	Ile	
1				5					10					15		
aca	gca	gaa	tta	atg	ccg	cca	aga	gga	gga	gac	ccc	gta	aga	tct	ctt	96
Thr	Ala	Glu	Leu	Met	Pro	Pro	Arg	Gly	Gly	Asp	Pro	Val	Arg	Ser	Leu	
			20					25						30		
aaa	ata	gca	caa	ctc	ttg	aga	aat	aag	gtg	cat	gca	ggt	aat	att	aca	144
Lys	Ile	Ala	Gln	Leu	Leu	Arg	Asn	Lys	Val	His	Ala	Val	Asn	Ile	Thr	
		35					40					45				
gac	gga	agt	aga	gca	ata	atg	aga	atg	tgt	agt	tta	gca	atg	tct	aaa	192
Asp	Gly	Ser	Arg	Ala	Ile	Met	Arg	Met	Cys	Ser	Leu	Ala	Met	Ser	Lys	
		50				55					60					
cta	tta	cta	gac	aat	ggg	ata	gaa	cct	ata	atg	cag	atc	tca	tgt	aga	240
Leu	Leu	Asp	Asn	Gly	Ile	Glu	Pro	Ile	Met	Gln	Ile	Ser	Cys	Arg		
				70				75						80		
gat	cgt	aat	aaa	att	gct	tta	caa	tca	gat	att	ctt	gga	gca	aat	gcc	288
Asp	Arg	Asn	Lys	Ile	Ala	Leu	Gln	Ser	Asp	Ile	Leu	Gly	Ala	Asn	Ala	
				85				90						95		
tta	gga	att	aaa	aat	att	tta	tgc	att	aca	gga	gat	tct	gta	aaa	gcc	336
Leu	Gly	Ile	Lys	Asn	Ile	Leu	Cys	Ile	Thr	Gly	Asp	Ser	Val	Lys	Ala	
			100					105						110		
gga	gat	cag	caa	gaa	aca	aaa	gcc	gtt	cat	gaa	ttt	gag	gca	gta	aga	384
Gly	Asp	Gln	Gln	Glu	Thr	Lys	Ala	Val	His	Glu	Phe	Glu	Ala	Val	Arg	
		115					120					125				
tta	tta	aaa	caa	att	caa	tca	ttc	aat	caa	gga	att	gat	cct	act	ttt	432
Leu	Leu	Lys	Gln	Ile	Gln	Ser	Phe	Asn	Gln	Gly	Ile	Asp	Pro	Thr	Phe	
		130				135					140					
gaa	caa	ctt	cca	gac	aaa	agg	act	gaa	att	ttc	tca	ggt	gcg	gca	gta	480
Glu	Gln	Leu	Pro	Asp	Lys	Arg	Thr	Glu	Ile	Phe	Ser	Gly	Ala	Ala	Val	
				150						155				160		
gat	cca	agt	tgt	cga	aat	caa	aga	agt	tta	aaa	agt	aga	aca	att	aaa	528
Asp	Pro	Ser	Cys	Arg	Asn	Gln	Arg	Ser	Leu	Lys	Ser	Arg	Thr	Ile	Lys	
				165				170						175		
aaa	aaa	gag	gcc	ggt	gca	aat	ttc	tta	caa	act	caa	ata	ggt	atg	gat	576
Lys	Lys	Glu	Ala	Gly	Ala	Asn	Phe	Leu	Gln	Thr	Gln	Ile	Val	Met	Asp	
			180					185					190			
aga	aaa	tgt	tta	gca	gac	ttt	tgc	aac	gaa	atc	agt	aat	cca	ctt	gag	624
Arg	Lys	Cys	Leu	Ala	Asp	Phe	Cys	Asn	Glu	Ile	Ser	Asn	Pro	Leu	Glu	
		195					200					205				
ata	cca	ggt	att	gca	gga	gta	ttt	ctt	tta	aaa	tca	tat	aaa	aat	gct	672
Ile	Pro	Val	Ile	Ala	Gly	Val	Phe	Leu	Leu	Lys	Ser	Tyr	Lys	Asn	Ala	
		210				215					220					
ctt	ttc	ata	aat	aaa	ttt	gta	cct	gga	gcg	aat	att	cct	gaa	aat	ggt	720
Leu	Phe	Ile	Asn	Lys	Phe	Val	Pro	Gly	Ala	Asn	Ile	Pro	Glu	Asn	Val	
225					230					235					240	

768

816

864

891

13> *Prochlorococcus maritima*

Arg Lys Cys Leu Ala Asp Phe Cys Asn Glu Ile Ser Asn Pro Leu Glu
195 200 205

Ile Pro Val Ile Ala Gly Val Phe Leu Leu Lys Ser Tyr Lys Asn Ala
 210 215 220

Leu Phe Ile Asn Lys Phe Val Pro Gly Ala Asn Ile Pro Glu Asn Val
 225 230 235 240

Leu Asn Arg Leu Lys Asp Ala Lys Asn Pro Leu Gln Glu Gly Ile Leu
 245 250 255

Ile Ala Ser Glu Gln Ala Gln Asp Phe Ile Asn Ile Ala Asp Gly Ile
 260 265 270

His Leu Met Ala Val Lys Ser Glu His Leu Ile Pro Glu Ile Leu Glu
 275 280 285

Lys Ala Gly Leu Asn Leu Glu Cys
 290 295

<210> 53
 <211> 1848
 <212> DNA
 <213> Bacillus stearothermophilus

<220>
 <221> CDS
 <222> (1) .. (1845)
 <223> RBE04103

<400> 53
 gtg gga ttg ctg gat gag ttg aaa gag cgc att ctc atc gcc gac ggg 48
 Val Gly Leu Leu Asp Glu Leu Lys Glu Arg Ile Leu Ile Ala Asp Gly
 1 5 10 15

gcg atg gga acg ctt tta tat tcg cac ggc att gac cgt tgt ttt gaa 96
 Ala Met Gly Thr Leu Leu Tyr Ser His Gly Ile Asp Arg Cys Phe Glu
 20 25 30

gaa ttg aat cta tcc aat cca gat gaa atc gtc cat att cat gaa gcg 144
 Glu Leu Asn Leu Ser Asn Pro Asp Glu Ile Val His Ile His Glu Ala
 35 40 45

atc gcc gcg ggc gcc gac gtc att cag acg aat aca tac ggc gcc 192
 Tyr Ile Ala Ala Gly Ala Asp Val Ile Gln Thr Asn Thr Tyr Gly Ala
 50 55 60

aac tat gtg aaa ctc gcc cgc tac ggc ctt gaa gat gag gtg ccg gcc 240
 Asn Tyr Val Lys Leu Ala Arg Tyr Gly Leu Glu Asp Glu Val Pro Ala
 65 70 75 80

atc aac cgc gcg gcg gtg cgg ctc gcc agg caa gcg gcg aac gga cgg 288
 Ile Asn Arg Ala Ala Val Arg Leu Ala Arg Gln Ala Ala Asn Gly Arg
 85 90 95

gca tac gtg ctc ggg acg atc ggg ggg ctg cgc acg tta aac aaa agc 336
 Ala Tyr Val Leu Gly Thr Ile Gly Gly Leu Arg Thr Leu Asn Lys Ser
 100 105 110

gtc gtc acg ctc gaa gaa gtg aag cgg acg ttt cgc gag cag ctg ttt 384
 Val Val Thr Leu Glu Glu Val Lys Arg Thr Phe Arg Glu Gln Leu Phe
 115 120 125

gtc ctg ctc gct gaa ggg gtc gac ggc gtg ctg ctc gag acg tat tac	432
Val Leu Leu Ala Glu Gly Val Asp Gly Val Leu Leu Glu Thr Tyr Tyr	
130 135 140	
gat ttg gaa gag ttg gag acg gtg ctt gcc atc gcc cgc aaa gag acc	480
Asp Leu Glu Glu Leu Thr Val Leu Ala Ile Ala Arg Lys Glu Thr	
145 150 155 160	
gac ttg ccg att atc gct cac gtc tcg ctc cat gaa gtc ggc gtc ttg	528
Asp Leu Pro Ile Ile Ala His Val Ser Leu His Glu Val Gly Val Leu	
165 170 175	
caa gat ggc acg ccg ctc gcg gac gcc ctt gcc cgc cta gag gcg ctc	576
Gln Asp Gly Thr Pro Leu Ala Asp Ala Leu Ala Arg Leu Glu Ala Leu	
180 185 190	
ggg gcc gat gtc gtc gga ctg aac tgt cgt ctc ggt cca tat cat atg	624
Gly Ala Asp Val Val Gly Leu Asn Cys Arg Leu Gly Pro Tyr His Met	
195 200 205	
ctt cgg tcg ctc gag gaa gtg ccg ctg cca aat cga gcg ttt ttg tcg	672
Arg Ser Leu Glu Glu Val Pro Leu Pro Asn Arg Ala Phe Leu Ser	
210 215 220	
gcg tat ccg aac gcc agc ctt ccg gat tac cgc gat ggg cgg ctt gtc	720
Ala Tyr Pro Asn Ala Ser Leu Pro Asp Tyr Arg Asp Gly Arg Leu Val	
225 230 235 240	
tat gag acg aac gct gaa tat ttc gag gaa acg gcc aaa gcg ttc cgc	768
Tyr Glu Thr Asn Ala Glu Tyr Phe Glu Glu Thr Ala Lys Ala Phe Arg	
245 250 255	
gac caa ggg gtg cgc ttg atc ggc ggg tgc tgc ggc acg acg ccg aaa	816
Asp Gln Gly Val Arg Leu Ile Gly Gly Cys Cys Gly Thr Thr Pro Lys	
260 265 270	
cat atc gaa gcg atg gca aaa gcg ctc tcc gac cga acg ccg gtg acg	864
His Ile Glu Ala Met Ala Lys Ala Leu Ser Asp Arg Thr Pro Val Thr	
275 280 285	
gaa aaa acg gtg aaa cgg cgc gcg gtg tct gta tca gtg caa gcg gag	912
Glu Lys Thr Val Lys Arg Arg Ala Val Ser Val Ser Val Gln Ala Glu	
290 295 300	
ggg ccc gcc cca tct ccc ctt ccc gag ctt gcc cgc acg cac cgc tcg	960
Arg Pro Ala Pro Ser Pro Leu Pro Glu Leu Ala Arg Thr His Arg Ser	
305 310 315 320	
gtc att gtg gag ctg gat ccg ccg aaa aaa ttg ggg att gac aag ttt	1008
Val Ile Val Glu Leu Asp Pro Pro Lys Lys Leu Gly Ile Asp Lys Phe	
325 330 335	
ctt gcc ggg gcg aaa gcg ctc cat gac gcc ggc atc gat gcg ctg acg	1056
Leu Ala Gly Ala Lys Ala Leu His Asp Ala Gly Ile Asp Ala Leu Thr	
340 345 350	
ttg gcc gac aac tcg ctc gcc acg ccg cgc atc agc aac gcc gct gtc	1104
Leu Ala Asp Asn Ser Leu Ala Thr Pro Arg Ile Ser Asn Ala Ala Val	
355 360 365	
gcc acg atc atc aag gag caa ctc ggc atc cgc ccg ctc gtg cat att	1152
Ala Thr Ile Ile Lys Glu Gln Leu Gly Ile Arg Pro Leu Val His Ile	
370 375 380	

aca tgc cgc gat cgc aat ttg atc ggc ttg cag tcg cat ttg atg ggc 1200
 Thr Cys Arg Asp Arg Asn Leu Ile Gly Leu Gln Ser His Leu Met Gly
 385 390 395 400
 ttg cat acg ctc ggc atc acc gat gtg ctc gcc att acc ggc gac ccg 1248
 Leu His Thr Leu Gly Ile Thr Asp Val Leu Ala Ile Thr Gly Asp Pro
 405 410 415
 tcg aaa atc ggc gat ttt cca ggg gca acg tcc gtg tac gac tta tca 1296
 Ser Lys Ile Gly Asp Phe Pro Gly Ala Thr Ser Val Tyr Asp Leu Ser
 420 425 430
 tcg ttc gat ttg atc cgc ttg atc cgc cag ttt aac gaa ggg ctg tcg 1344
 Ser Phe Asp Leu Ile Arg Leu Ile Arg Gln Phe Asn Glu Gly Leu Ser
 435 440 445
 tac tcg ggc aaa ccg ctt ggg caa aaa acg aac ttc tcg atc ggc gct 1392
 Tyr Ser Gly Lys Pro Leu Gly Gln Lys Thr Asn Phe Ser Ile Gly Ala
 450 455 460
 g ttc aac ccg aac gtc cgc cat ttg gac aaa gcg gtc gag cgg atg 1440
 Phe Asn Pro Asn Val Arg His Leu Asp Lys Ala Val Glu Arg Met
 470 475 480
 gag aaa aaa atc caa tgc ggc gcc cat tat ttc ttg acc cag ccg att 1488
 Glu Lys Lys Ile Gln Cys Gly Ala His Tyr Phe Leu Thr Gln Pro Ile
 485 490 495
 tac tcg gaa gag aaa atc gtt gaa gtg cac gaa gcg acc aag cat ctt 1536
 Tyr Ser Glu Glu Lys Ile Val Glu Val His Glu Ala Thr Lys His Leu
 500 505 510
 gac acg ccg att tac atc ggc att atg ccg ctt gtg agc gcg cgc aac 1584
 Asp Thr Pro Ile Tyr Ile Gly Ile Met Pro Leu Val Ser Ala Arg Asn
 515 520 525
 gcc gac ttt ttg cat cat gaa gtg ccg ggc att acg ctc tct gac gag 1632
 Ala Asp Phe Leu His His Glu Val Pro Gly Ile Thr Leu Ser Asp Glu
 530 535 540
 att cgc gcc cgc atg gcc gcc tgc agc ggc gac ccg gtg caa gca gcc 1680
 Arg Ala Arg Met Ala Ala Cys Ser Gly Asp Pro Val Gln Ala Ala
 550 555 560
 aag gaa ggc atc gct atc gcc aaa tcg ctc att gac gct gcg ttt gat 1728
 Lys Glu Gly Ile Ala Ile Ala Lys Ser Leu Ile Asp Ala Ala Phe Asp
 565 570 575
 ttg ttt aac ggc att tat ttg atc acg ccg ttc ttg cgc tac gac atg 1776
 Leu Phe Asn Gly Ile Tyr Leu Ile Thr Pro Phe Leu Arg Tyr Asp Met
 580 585 590
 acg gtc gag ctt gtc cgc tac att cac gaa aaa gaa gcg gcc gcc aaa 1824
 Thr Val Glu Leu Val Arg Tyr Ile His Glu Lys Glu Ala Ala Ala Lys
 595 600 605
 gaa agg aag gtt gtt cat ggc taa 1848
 Glu Arg Lys Val Val His Gly
 610 615

<210> 54

M/43126

MetF

<211> 615

<212> PRT

<213> *Bacillus stearothermophilus*

<400> 54

Val	Gly	Leu	Leu	Asp	Glu	Leu	Lys	Glu	Arg	Ile	Leu	Ile	Ala	Asp	Gly
1				5					10					15	

Ala	Met	Gly	Thr	Leu	Leu	Tyr	Ser	His	Gly	Ile	Asp	Arg	Cys	Phe	Glu
			20					25					30		

Glu	Leu	Asn	Leu	Ser	Asn	Pro	Asp	Glu	Ile	Val	His	Ile	His	Glu	Ala
		35					40					45			

Tyr	Ile	Ala	Ala	Gly	Ala	Asp	Val	Ile	Gln	Thr	Asn	Thr	Tyr	Gly	Ala
		50				55					60				

Asn	Tyr	Val	Lys	Leu	Ala	Arg	Tyr	Gly	Leu	Glu	Asp	Glu	Val	Pro	Ala
65					70					75					80

Asn	Arg	Ala	Ala	Val	Arg	Leu	Ala	Arg	Gln	Ala	Ala	Asn	Gly	Arg	
			85					90					95		

Ala	Tyr	Val	Leu	Gly	Thr	Ile	Gly	Gly	Leu	Arg	Thr	Leu	Asn	Lys	Ser
			100					105					110		

Val	Val	Thr	Leu	Glu	Glu	Val	Lys	Arg	Thr	Phe	Arg	Glu	Gln	Leu	Phe
		115					120					125			

Val	Leu	Leu	Ala	Glu	Gly	Val	Asp	Gly	Val	Leu	Leu	Glu	Thr	Tyr	Tyr
		130				135						140			

Asp	Leu	Glu	Glu	Leu	Glu	Thr	Val	Leu	Ala	Ile	Ala	Arg	Lys	Glu	Thr
145					150					155					160

Asp	Leu	Pro	Ile	Ile	Ala	His	Val	Ser	Leu	His	Glu	Val	Gly	Val	Leu
			165						170					175	

Gln	Asp	Gly	Thr	Pro	Leu	Ala	Asp	Ala	Leu	Ala	Arg	Leu	Glu	Ala	Leu
		180					185						190		

Ala	Asp	Val	Val	Gly	Leu	Asn	Cys	Arg	Leu	Gly	Pro	Tyr	His	Met	
	195					200					205				

Leu	Arg	Ser	Leu	Glu	Glu	Val	Pro	Leu	Pro	Asn	Arg	Ala	Phe	Leu	Ser
	210					215					220				

Ala	Tyr	Pro	Asn	Ala	Ser	Leu	Pro	Asp	Tyr	Arg	Asp	Gly	Arg	Leu	Val
225					230					235					240

Tyr	Glu	Thr	Asn	Ala	Glu	Tyr	Phe	Glu	Glu	Thr	Ala	Lys	Ala	Phe	Arg
			245						250					255	

Asp	Gln	Gly	Val	Arg	Leu	Ile	Gly	Gly	Cys	Cys	Gly	Thr	Thr	Pro	Lys
		260					265					270			

His	Ile	Glu	Ala	Met	Ala	Lys	Ala	Leu	Ser	Asp	Arg	Thr	Pro	Val	Thr
		275				280					285				

Glu	Lys	Thr	Val	Lys	Arg	Arg	Ala	Val	Ser	Val	Ser	Val	Gln	Ala	Glu
	290					295					300				

Arg Pro Ala Pro Ser Pro Leu Pro Glu Leu Ala Arg Thr His Arg Ser
 305 310 315 320
 Val Ile Val Glu Leu Asp Pro Pro Lys Lys Leu Gly Ile Asp Lys Phe
 325 330 335
 Leu Ala Gly Ala Lys Ala Leu His Asp Ala Gly Ile Asp Ala Leu Thr
 340 345 350
 Leu Ala Asp Asn Ser Leu Ala Thr Pro Arg Ile Ser Asn Ala Ala Val
 355 360 365
 Ala Thr Ile Ile Lys Glu Gln Leu Gly Ile Arg Pro Leu Val His Ile
 370 375 380
 Thr Cys Arg Asp Arg Asn Leu Ile Gly Leu Gln Ser His Leu Met Gly
 385 390 395 400
 Leu His Thr Leu Gly Ile Thr Asp Val Leu Ala Ile Thr Gly Asp Pro
 405 410 415
 Lys Ile Gly Asp Phe Pro Gly Ala Thr Ser Val Tyr Asp Leu Ser
 420 425 430
 Ser Phe Asp Leu Ile Arg Leu Ile Arg Gln Phe Asn Glu Gly Leu Ser
 435 440 445
 Tyr Ser Gly Lys Pro Leu Gly Gln Lys Thr Asn Phe Ser Ile Gly Ala
 450 455 460
 Ala Phe Asn Pro Asn Val Arg His Leu Asp Lys Ala Val Glu Arg Met
 465 470 475 480
 Glu Lys Lys Ile Gln Cys Gly Ala His Tyr Phe Leu Thr Gln Pro Ile
 485 490 495
 Tyr Ser Glu Glu Lys Ile Val Glu Val His Glu Ala Thr Lys His Leu
 500 505 510
 Asp Thr Pro Ile Tyr Ile Gly Ile Met Pro Leu Val Ser Ala Arg Asn
 515 520 525
 Asp Phe Leu His His Glu Val Pro Gly Ile Thr Leu Ser Asp Glu
 530 535 540
 Ile Arg Ala Arg Met Ala Ala Cys Ser Gly Asp Pro Val Gln Ala Ala
 545 550 555 560
 Lys Glu Gly Ile Ala Ile Ala Lys Ser Leu Ile Asp Ala Ala Phe Asp
 565 570 575
 Leu Phe Asn Gly Ile Tyr Leu Ile Thr Pro Phe Leu Arg Tyr Asp Met
 580 585 590
 Thr Val Glu Leu Val Arg Tyr Ile His Glu Lys Glu Ala Ala Ala Lys
 595 600 605
 Glu Arg Lys Val Val His Gly
 610 615

<210> 55

<211> 52

<212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz:PCR primer

 <400> 55
 cccgggatcc gctagcggcg cgccggccgg cccggtgtga aataccgcac ag 52

 <210> 56
 <211> 53
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz:PCR primer

 <400> 56
 tctagactcg agcggccgcg gccggccttt aaattgaaga cgaaagggcc tcg 53

 <210> 57
 <211> 47
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz:PCR primer

 <400> 57
 gagatctaga cccggggatc cgctagcggg ctgctaaagg aagcgga 47

 <210> 58
 <211> 38
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz:PCR primer

 <400> 58
 gagggcgcg ccgctagcgt gggcgaagaa ctccagca 38

 <210> 59
 <211> 34
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz:PCR primer

 <400> 59
 gagagggcgg ccgcgcaaag tcccgttcg tgaa 34

 <210> 60
 <211> 34
 <212> DNA
 <213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 60

gagagggcgg ccgctcaagt cggtaagcc acgc

34

<210> 61

<211> 140

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 61

tcgaatttaa atctcgagag gcctgacgtc gggcccggta ccacgcgtca tatgactagt 60
 tcggacctag ggatatcgtc gacatcgatg ctcttctgcg ttaattaaca attgggatcc 120
 tctagaccg ggattttaa 140

<210> 62

<211> 140

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 62

gatcatttaa atcccggggtc tagaggatcc caattgttaa ttaacgcaga agagcatcga 60
 tgtcgacgat atccctaggt ccgaactagt catatgacgc gtggtaccgg gcccgacgtc 120
 aggctctcg agattttaa 140

<210> 63

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 63

gagagcggcc gccgatcctt tttaacccat cac

33

<210> 64

<211> 32

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 64

aggagcggcc gccatcggca ttttcttttg cg

32

<210> 65

<211> 5091

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Plasmid

<400> 65

```

gccgcgactg ccttcgcgaa gccttgcccc gcggaaattt cctccaccga gttcgtgcac 60
acccttatgc caagcttctt tcaccctaaa ttcgagagat tggattctta ccgtggaaat 120
tcttcgcaaa aatcgtcccc tgatcgccct tgcgacgttg gcgtcggtgc cgctgggtgc 180
gcttggtctg accgacttga tcagcgcccg ctcgatttaa atctcgagag gctgacgtc 240
gggcccggta ccacgcgtca tatgactagt tcggacctag ggatatcgtc gacatcgatg 300
ctcttctgcg ttaattaaca attgggatcc tctagaccgg ggatttaaat cgctagcggg 360
ctgctaaagg aagcggaaca cgtagaaagc cagtccgcag aaacgggtgct gaccccggat 420
gaatgtcagc tactgggcta tctggacaag ggaaacgcga agcgcaaaga gaaagcaggt 480
agcttgcaag gggcttacat ggcgatagct agactgggcg gttttatgga cagcaagcga 540
accggaattg ccagctgggg cgccctctgg taaggttggg aagccctgca aagtaaactg 600
gatggctttc ttgcccgaag ggatctgagt gcgcagggga tcaagatctg atcaagagac 660
aggatgagga tcgtttcgca tgattgaaca agatggattg cacgcaggtt ctccggccgc 720
ttgggtggag aggctattcg gctatgactg ggcacaacag acaatcggct gctctgatg 780
cgccgtgttc cggctgtcag cgcaggggcg cccggttctt tttgtcaaga ccgacctgtc 840
cgggtgccctg aatgaactgc aggacgaggg agcgcggtta tcgtggctgg ccacgacggg 900
cgttccttgc gcagctgtgc tcgacgttgt cactgaagcg ggaagggact ggctgctatt 960
gcgaagtg ccggggcagg atctcctgtc atctcacctt gctcctgccc agaaagtatc 1020
gatggctc gatgcaatgc ggcggtgca tacgcttgat ccggctacct gccattcga 1080
ccaagcg aaacatcgca tcgagcgagc acgtactcgg atggaagccg gtcttgtcga 1140
tcaggatgat ctggacgaag agcatcaggg gctcgcgcca gccgaactgt tcgccaggct 1200
caaggcgcg atgcccgcg gcgaggatct cgtcgtgacc catggcgatg cctgcttgcc 1260
gaatatcatg gtggaaaatg gccgcttttc tggattcatc gactgtggcc ggctgggtgt 1320
ggcggaaccg tatcaggaca tagcgttggc taccgctgat attgctgaag agcttggcgg 1380
cgaatgggct gaccgcttcc tcgtgcttta cggtatcgcc gctcccgat cgacgcgat 1440
cgcttcttat cgcttcttgc acgagttctt ctgagcggga ctctggggtt cgaaatgacc 1500
gaccaagcga cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa 1560
aggttgggct tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat 1620
ctcatgctgg agttcttcgc ccacgctagc ggcgcgcccg ccggcccggg gtgaaatacc 1680
gcacagatgc tgggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 1800
ctcgtgcgc tcggtcgttc gggataacgc aggaagaac atgtgagcaa aaggccagca 1860
acgggttatc aaacgtaaaa aggcgcgctt gctggcggtt ttccataggc tccgcccccc 1920
tgacgagcat caaaaaaatc gacgctcaag tcagaggtgg cgaaaccoga caggactata 1980
aagataccga gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgacctgcc 2040
gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc 2100
acgctgtagg tatctcagtt cgtgttaggt gctgcgcctt atccggtaac tatcgtcttg agtccaacct 2220
acccccggt taagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 2280
atgttaggc ggtgctacag agttcttgaa gtgggtggcct aactacggct acactagaag 2340
gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 2400
ctcttgatcc ggcaaaacaa ccaccgctgg tagcgggtgg ttttttgttt gcaagcagca 2460
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 2520
cgctcagtgg aacgaaaact cacgttaagg gattttgggtc atgagattat caaaaaggat 2580
cttcacctag atccttttaa aggcgggccc cggccgcgca aagtcccgtc tcgtgaaaat 2640
tttcgtgccc cgtgattttc cgccaaaaac tttaacgaac gttcgttata atggtgtcat 2700
gaccttcacg acgaagtact aaaattggcc cgaatcatca gctatggatc tctctgatgt 2760
cgcgctggag tccgacgcgc tcgatgctgc cgtcgattta gctatggatc tctctgatgt 2760
ccgagctctc gatagcagcg acgcgcagc atcacgagac tgggccagtg ccgcgagcga 2880
cctagaaact ctggtggcgg atcttgagga gctggctgac gagctgcgtg ctccggccagc 2940
gocaggagga cgcacagtag tggaggatgc aatcagttgc gcctactgcg gtggcctgat 3000
tcttccccgg cctgaccgcg gaggacggcg cgcaaaatat tgctcagatg cgtgtcgtgc 3060
cgagccagc cgcgagcgcg ccaacaaacg ccacgccgag gagctggagg cggctaggtc 3120
gcaaatggcg ctggaaagtgc gtcccccgag cgaaattttg gccatggtcg tcacagagct 3180
ggaagcggca gcgagaatta tcgcgatcgt ggcgggtgcc gcaggcatga caaacatcgt 3240
aaatgccgog tttcgtgtgc cgtggccgcc caggagctgt cagcgccgcc accacctgca 3300
ccgaatcggc agcagcgtcg cgcgtcgaag aagcgcacag gcgcaagaa gcgataagct 3360
gcacgaatac ctgaaaaatg ttgaacgcc cgtgagcggt aactcacagg gcgtcggcta 3420
acccccagtc caaacctggg agaaagcgct caaaaatgac tctagcggat tcacgagaca 3480

```

```

ttgacacacc ggcctggaaa ttttccgctg atctgttctga caccatccc gagctcgcgc 3540
tgcatcacg tggctggacg agcgaagacc gccgcgaatt cctcgtcac ctgggcagag 3600
aaaatttcca gggcagcaag acccgcgact tcgccagcgc ttggatcaaa gacccggaca 3660
cggagaaaca cagccgaagt tataccgagt tggttcaaaa tcgcttgccc ggtgccagta 3720
tgttgctctg acgcacgcgc agcacgcagc cgtgcttgctc ctggacattg atgtgccgag 3780
ccaccaggcc ggcgggaaaa tcgagcacgt aaaccccag gtctacgca ttttggagcg 3840
ctgggcacgc ctggaaaaag cgcagcttg gatcggcgtg aatccactga gcgggaaatg 3900
ccagctcatc tggctcattg atccggtgta tgccgcagca ggcattgagca gccgaatat 3960
gcgcctgctg gctgcaacga ccgaggaaat gacccgcgtt ttcggcgtg accaggcttt 4020
ttcacatagg ctgagccgtg gccactgcac tctccgacga tcccagccgt accgctggca 4080
tgcccagcac aatcgcgtgg atcgcttagc tgatcttatg gaggttgctc gcatgatctc 4140
aggcacagaa aaacctaaaa aacgctatga gcaggagttt tctagcggac gggcacgtat 4200
cgaagcggca agaaaagcca ctgcggaagc aaaagcactt gccacgcttg aagcaagcct 4260
gccgagcgc gctgaagcgt ctggagagct gatcgacggc gtccgtgtcc tctggactgc 4320
tccaggcgt gccgcccgtg atgagacggc ttttcgccac gctttgactg tgggatacca 4380
gttaaaagcg gctggtgagc gcctaaaaa cccaagggt catcgagcct acgagcgtgc 4440
ctacaccgtc gctcaggcgg tcggaggagg cgtgagcct gatctgccgc cggactgtga 4500
ccgccagacg gattggcgcg gacgtgtgcg cggctacgtc gctaaaggcc agccagtcgt 4560
ccctgctcgt cagacagaga cgcagagcca gccgaggcga aaagctctgg ccactatggg 4620
aagacgtggc ggtaaaaagg ccgcagaacg ctggaaagac ccaaacagtg agtacgcccg 4680
agcacagcga gaaaaactag ctaagtccag tcaacgacaa gctaggaaag ctaaaggaaa 4740
gcttgacc attgcaggtt ggtttatgac tgttgaggga gagactggct cgtggccgac 4800
caatgaa gctatgtctg aatttagcgt gtcacgtcag accgtgaata gagcacttaa 4860
ctgcggg cattgaactt ccacgaggac gccgaaagct tcccagtaaa tgtgccatct 4920
cgtaggcaga aaacggttcc cccgtagggt ctctctcttg gcctcctttc taggtcgggc 4980
tgattgctct tgaagctctc taggggggt cacaccatag gcagataacg ttccccaccg 5040
gctcgctcgt taagcgcaca aggactgctc ccaaagatct tcaaagccac t 5091

```

<210> 66

<211> 4323

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Plasmid

<400> 66

```

tctctcagcg tatggttgct gcctgagctg tagttgcctt catcgatgaa ctgctgtaca 60
ttttgatacg ttttccgctc accgtcaaag attgatttat aatcctctac accgttgatg 120
ttcaaagagc tgtctgatgc tgatacgta actgtgcag ttgtcagtg ttgtttgccg 180
atgtttac cggagaaatc agtgtagaat aaacggattt ttccgtcaga tgtaaatgtg 240
tgaacctg accattcttg tgtttggtct tttaggatag aatcatttgc atcgaatttg 300
tcgctgtctt taaagacgcg gccagcgttt ttccagctgt caatagaagt ttcgccgact 360
ttttgataga acatgtaaat cgatgtgtca tccgcatttt taggatctcc ggctaatagca 420
aagacgatgt ggtagccgtg atagtttgcg acagtgccgt cagcgttttg taatggccag 480
ctgtcccaaa cgtccaggcc ttttgcagaa gagatatttt taattgtgga cgaatcaaat 540
tcagaaaactt gatatttttc atttttttgc tgttcaggga tttgcagcat atcatggcgt 600
gtaatatggg aaatgccgta tgtttcctta tatggctttt ggttcgtttc tttcgcaaac 660
gcttgagttg cgcctcctgc cagcagtgcg gtagttaaag ttaatactgt tgcttgtttt 720
gcaaaacttt tgatgttcat cgttcatgtc tcttttttta tgtactgtgt tagcggctctg 780
cttcttccag cctcctgttt tgaagatggc aagttagtta cgcacaataa aaaaagacct 840
aaaatatgta aggggtgacg ccaaagtata cactttgccc tttacacatt ttaggtcttg 900
cctgctttat cagtaacaaa cccgcgcgat ttacttttgc acctcattct attagactct 960
cgtttgatt gcaactggtc tattttcttc ttttgtttga tagaaaatca taaaaggatt 1020
tgcagactac gggcctaaag aactaaaaaa tctatctggt tcttttcatt ctctgtattt 1080
tttatagttt ctggtgcatg ggcataaagt tgccttttta atcacaattc agaaaatatc 1140
ataatatctc atttcaacta ataatagtga acggcaggta tatgtgatgg gttaaaaagg 1200
atcggcggcc gctcgattta aatctcgaga ggcctgacgt cgggcccggg accacgcgtc 1260
atatgactag ttcggacctc gggatatcgt cgacatcgat gctcttctgc gttaattaac 1320
aattgggac cctagacccc gggattttaa tcgctagcgg gctgctaaag gaagcggaac 1380
acgtagaaag ccagtcgcga gaaacggtgc tgaccccgga tgaatgtcag ctactgggct 1440

```

M/43126

MetF

atctggacaa	gggaaaacgc	aagcgcaaag	agaaagcagg	tagcttgacg	tgggcttaca	1500
tggcgatagc	tagactgggc	ggttttatgg	acagcaagcg	aaccggaatt	gccagctggg	1560
gcgccctctg	gtaagggttg	gaagccctgc	aaagtaaact	ggatggcttt	cttgccgcca	1620
aggatctgat	ggcgagggg	atcaagatct	gatcaagaga	caggatgagg	atcgtttcgc	1680
atgattgaac	aagatggatt	gcacgcaggt	tctccggccg	cttgggtgga	gaggctattc	1740
ggctatgact	gggcacaaca	gacaatcggc	tgctctgatg	ccgccgtgtt	ccggctgtca	1800
gcgcaggggc	gcccggttct	ttttgtcaag	accgacctgt	ccggtgccct	gaatgaactg	1860
caggacgagg	cagcgcggtc	atcgtggctg	gccacgacgg	gcgttccttg	cgcagctgtg	1920
ctcgacgttg	tactgaagc	gggaagggac	tggctgctat	tgggcgaagt	gccggggcag	1980
gatctcctgt	catctcacct	tgctcctgcc	gagaaagtat	ccatcatggc	tgatgcaatg	2040
cggcggtgtc	atacgcttga	tccggctacc	tgcccatctg	accaccaagc	gaaacatcgc	2100
atcgagcgag	cacgtactcg	gatggaagcc	ggtcttgctg	atcaggatga	tctggacgaa	2160
gagcatcagg	ggctcgcgcc	agccgaactg	tccgccaggc	tcaaggcgcg	catgcccgcg	2220
ggcgaggatc	tcgtcgtgac	ccatggcgat	gcctgcttgc	cgaatatcat	ggtggaaaat	2280
ggccgctttt	ctggattcat	cgactgtggc	cggctgggtg	tggcggaaccg	ctatcaggac	2340
atagcgttgg	ctaccogtga	tattgctgaa	gagcttggcg	gcgaatgggc	tgaccgcttc	2400
ctcgtgcttt	acgggtatcg	cgctcccgat	tcgcagcgca	tcgccttcta	tcgccttctt	2460
gacgagttct	tctgagcggg	actctgggtt	tcgaaatgac	cgaccaagcg	acgcccgaac	2520
tgccatcacg	agattttgat	tccaccgccg	ccttctatga	aaggttgggc	ttcggaatcg	2580
ttttccggga	cgccggctgg	atgatcctcc	agcgcgggga	tctcatgctg	gagttcttcg	2640
cccacgctag	cggcgcgccg	gccggcccgg	tgtgaaatac	cgcacagatg	cgtaaggaga	2700
ctaccgca	tcaggcgctc	ttccgcttcc	tcgctcactg	actcgctgcg	ctcggtcggt	2760
ctgcggc	gagcggtatc	agctcactca	aaggcggtaa	tacggttatc	cacagaatca	2820
tgataacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	gaaccgtaaa	2880
aaggccgcgt	tgctggcggt	tttccatagg	ctccgcccc	ctgacgagca	tcacaaaaat	2940
cgacgctcaa	gtcagaggtg	gcgaaaccgg	acaggactat	aaagatacca	ggcgtttccc	3000
cctggaagct	ccctcgtgcg	ctctcctggt	ccgaacctgc	cgcttaccgg	atacctgtcc	3060
gcctttctcc	cttcgggaag	cgtggcgctt	tctcatagct	cacgctgtag	gtatctcagt	3120
tcggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	aacccccgt	tcagcccgcg	3180
cgctgcgcct	tatccggtaa	ctatcgtcct	gagtccaacc	cggtaagaca	cgacttatcg	3240
ccactggcag	cagccactgg	taacaggatt	agcagagcga	ggtatgtagg	cggtgctaca	3300
gagttcttga	agtgggtggc	taactacggc	tacactagaa	ggacagtatt	tggtatctgc	3360
gctctgctga	agccagttac	cttcggaaaa	agagttggta	gctcttgatc	cggcaaaaca	3420
accaccgctg	gtagcggttg	tttttttggt	tgcaagcagc	agattacgcg	cagaaaaaaa	3480
ggatctcaag	aagatccttt	gatcttttct	acggggtctg	acgctcagtg	gaacgaaaac	3540
tcacgttaag	ggatttttgg	catgagatta	tcaaaaagga	tcttcacctc	gatcctttta	3600
aaggccggcc	gcggccgcca	tcggcatttt	cttttgctgt	tttatttggt	aactgttaat	3660
tgtccttggt	caaggatgct	gtctttgaca	acagatgttt	tcttgccctt	gatgttcagc	3720
aggaagctcg	gcgcaaactg	tgattgtttg	tctgcgtaga	atcctctggt	tgtcatatag	3780
cttgtaatca	cgacattggt	tccttttcgct	tgaggtacag	cgaagtgtga	gtaagtaaag	3840
gttacatcgt	taggatcaag	atccattttt	aacacaaggc	cagttttggt	cagcggcttg	3900
ctggggccag	ttaaagaatt	agaaacataa	ccaagcatgt	aaatatcggt	agacgtaatg	3960
ctcaatcg	tcatttttga	tccgcgggag	tcagtgaaca	ggtaccattt	gccgttcatt	4020
ctaaagacgt	tcgcgcgttc	aattttcatct	gttactgtgt	tagatgcaat	cagcggtttc	4080
atcacttttt	tcagtgtgta	atcatcgctt	agctcaatca	taccgagagc	gccgtttgct	4140
aactcagccg	tcggtttttt	atcgctttgc	agaagttttt	gactttcttg	acggaagaat	4200
gatgtgcttt	tgccatagta	tgctttgtta	aataaagatt	cttcgccttg	gtagccatct	4260
tcagttccag	tgttttgcttc	aaataactaag	tatttgtggc	ctttatcttc	tacgtagtga	4320
gga						4323